

Existence condition of strong stationary times for continuous time Markov chains on discrete graphs

Guillaume Copros

July 2016

Abstract

We consider a random walk on a discrete connected graph having some infinite branches plus finitely many vertices with finite degrees. We find the generator of a strong stationary dual in the sense of Fill, and use it to find some equivalent condition to the existence of a strong stationary time. This strong stationary dual process lies in the set of connected compact sets of the compactification of the graph. When this graph is \mathbb{Z} , this is simply the set of (possibly infinite) segments of \mathbb{Z} .

Keywords : Strong stationary time, Strong stationary dual, Random walk, Discrete graph

MSC2010 : 60J27, 60G40

1 Introduction

1.1 Historical background

Convergence of Markov processes to their stationary distribution has been much studied by analytical tools. For instance, classical theorems provide ways to find a deterministic time t_0 at which the distribution of a process X will be arbitrarily close to its stationary distribution μ_∞ , for the total variation distance.

A more probabilistic approach is to look for a random time $T < \infty$ such that X_T is distributed *exactly* as its stationary distribution, independently of T . Such a T is called a *strong stationary time*, and will be defined more precisely in (3). This tool was first introduced and developed by Aldous and Diaconis [1] for discrete time Markov chains with finite state space. In this framework, Diaconis and Fill [3] then introduced a method to construct strong stationary times by intertwining processes. This work was transposed to continuous time Markov chains by Fill [7] and [6].

More recently, strong stationary times have been investigated, among others, by Lyzinski and Fill [5], Miclo [12] Gong, Mao and Zhang [10] in continuous time, and Lorek and Szekli [11] in discrete time.

They have also been used by Diaconis and Saloff-Coste [4] to study cut-off phenomenon, Fill [8] for perfect sampling or Fill and Kahn [9] for fastest mixing.

A question that naturally arises is: does it always exist a finite strong stationary time? On a finite state space, Diaconis and Fill [3] proved that the answer is yes. Unfortunately, this is no longer true when the state space is infinite, even countable. Of course, the existence or not depends on the initial distribution: when it is the same as the stationary one, 0 is an obvious strong stationary time! One can therefore wonder under which conditions there exists a finite one, for a given process, whatever the initial distribution. Thus, Miclo [12] produced some existence conditions in the case of a diffusion on \mathbb{R} . We aim here to give similar conditions for continuous time random walks on a connected graph where all vertices but a finite number have degree two.

In Section 2 we adapt an important result from Fill [7] to our case, that we will apply in Section 3 to birth-death processes on \mathbb{Z} , and in Section 4 to random walks on a graph.

1.2 Some reminders on Markov processes

What follows recalls briefly the characterization of continuous time Markov chains in terms of generators and semi-groups. For more details, we refer to Norris [14]. The stochastic processes will be defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, implied in the sequel, which we assume to be "large enough". For a Markov process $(X_t)_{t \geq 0}$ on $(\Omega, \mathcal{A}, \mathbb{P})$, we will denote by $X_{[0,t]}$ its trajectory until the time t ($t \geq 0$), $(\mathcal{F}_t^X)_{t \geq 0}$ the filtration generated by X and \mathcal{F}_∞^X the σ -algebra generated by the union of \mathcal{F}_t^X . If τ is a (\mathcal{F}_t^X) -stopping time, we write:

$$(\mathcal{F}_\tau^X) := \{A \in \mathcal{F}_\infty^X : A \cap \{\tau \leq t\} \in \mathcal{F}_t^X, \forall t \geq 0\}$$

A random time T is called a *randomized stopping time* relative to X if there exists a sub σ -algebra \mathcal{G} of \mathcal{A} , independent from \mathcal{F}_∞^X , such that T is a stopping time with respect to $(\sigma(\mathcal{F}_t^X, \mathcal{G}))_{t \geq 0}$ (cf Fill [6], Section 2.2).

Let \mathcal{S} be a countable topological space (not necessarily discrete). We add to this space an isolated point Δ , called *cemetery point*, and we extend every real function f on \mathcal{S} by setting $f(\Delta) = 0$. A *stable and conservative Q-matrix* on \mathcal{S} is a matrix $(\mathcal{L}_{x,y})_{(x,y) \in \mathcal{S}^2}$ satisfying:

$$0 \leq \mathcal{L}_{x,y} < +\infty, \quad \forall x \neq y \in \mathcal{S}$$

$$\sum_{y \neq x} \mathcal{L}_{x,y} = -\mathcal{L}_{x,x} < +\infty, \quad \forall x \in \mathcal{S}$$

For every Q-matrix \mathcal{L} , we will write $\mathcal{L}_x := -\mathcal{L}_{x,x}$, and we enlarge \mathcal{L} to $\mathcal{S} \cup \{\Delta\}$ by setting $\mathcal{L}_{x,\Delta} = \mathcal{L}_{\Delta,x} = \mathcal{L}_{\Delta,\Delta} = 0$ for every $x \in \mathcal{S}$. Together with a probability measure μ_0 on \mathcal{S} , a Q-matrix \mathcal{L} defines a process X with values in $\mathcal{S} \cup \{\Delta\}$, through its jump chain $(Y_n)_{n \in \mathbb{N}}$ and jumping times $(T_n)_{n \in \mathbb{N}}$:

- $(Y_n)_{n \in \mathbb{N}}$ is a discrete time Markov chain, with initial distribution μ_0 and transition matrix $\left((1 - \delta_{x,y}) \frac{\mathcal{L}_{x,y}}{\mathcal{L}_x} \right)_{(x,y) \in \mathcal{S}^2}$, where $\delta_{x,y}$ is the Kronecker delta, equal to 1 if $x = y$ and 0 otherwise.
- $T_0 = 0$, and for every $n \in \mathbb{N}^*$, given Y_0, \dots, Y_n , the random variables $T_1 - T_0, \dots, T_n - T_{n-1}$ are independent and exponentially distributed with respective parameters $\mathcal{L}_{Y_0}, \dots, \mathcal{L}_{Y_{n-1}}$.
- For all $i \in \mathbb{N}$ and all $t \in [T_i, T_{i+1}[$, we set $X_t = Y_i$.
- Define $T := \lim_{n \rightarrow \infty} T_n$. If $T < \infty$, we set $X_t = \Delta$ for every $t \geq T$.

This process is Markovian, homogeneous, and its trajectories are right continuous with left limits (càdlàg). In addition, this construction uniquely determines the law of X . In the following, we will simply say that \mathcal{L} is a generator on \mathcal{S} , and that X is a *minimal process* with generator \mathcal{L} and initial distribution μ_0 , or only with generator \mathcal{L} whenever the initial distribution does not matter. The time T is called the *explosion time* of X , and we will see in Section 2.2 how in certain cases we can naturally construct a non-minimal process, that is, one that is not "killed" after the explosion time. In the case where $T = \infty$ almost surely for every initial distribution, the generator \mathcal{L} (or the process X) is said to be nonexplosive, and explosive otherwise. It is positive recurrent (respectively irreducible, reversible) if the jump chain $(Y_n)_{n \in \mathbb{N}}$ is so, for every initial distribution.

The law of such a process is also determined by its finite dimensional distributions, namely the set of

$$\mathbb{P}(X_{t_1} = x_1, \dots, X_{t_n} = x_n)$$

with $n \in \mathbb{N}$, $0 \leq t_1 < \dots < t_n < \infty$ and $x_1, \dots, x_n \in \mathcal{S}$. The *transition function* of X is defined by:

$$P_{x,y}(t) = \mathbb{P}(X_t = y \mid X_0 = x)$$

and is then (together with an initial distribution μ_0) another way to characterize the law of X , through its finite dimensional distributions. A transition function defines a *semi-group* $(P(t))_{t \geq 0}$ of sub-stochastic matrices, which may be enlarged to $\mathcal{S} \cup \{\Delta\}$ to become stochastic.

A semi-group $P(\cdot)$ is a *solution of a Kolmogorov equation* associated to \mathcal{L} if for every $t \geq 0$ one of the matrix equalities:

$$P'(t) = \mathcal{L}P(t) \quad (\text{backward equation}) \quad (1)$$

$$P'(t) = P(t)\mathcal{L} \quad (\text{forward equation}) \quad (2)$$

hold. It is the minimal solution of these equations if for every other solution $\tilde{P}(\cdot)$ we have $P_{x,y}(t) \leq \tilde{P}_{x,y}(t)$, for every $x, y \in \mathcal{S}$ and every $t \geq 0$. We recall that the minimal solution exists (theorem 2.8.3 and 2.8.6 of [14]), and is necessarily unique by definition.

Each of the three following conditions is implied by the other two:

- X is a minimal process with generator \mathcal{L}
- X is a minimal process with semi-group $P(\cdot)$
- $P(\cdot)$ is the minimal solution of the backward (respectively forward) equation associated to \mathcal{L}

Finally, although we will only use the characterization of processes through Q -matrices and transition functions, we also recall a useful consequence in the functional point of view. Let $\mathcal{B}(\mathcal{S})$ be the set of real valued, bounded functions on \mathcal{S} . We see a function $f \in \mathcal{B}(\mathcal{S})$ as a column vector indexed by \mathcal{S} , and in that case $P(\cdot)$ and \mathcal{L} act on $\mathcal{B}(\mathcal{S})$ by matrix multiplication.

If X is a minimal process with generator \mathcal{L} , and τ_n , $n \in \mathbb{N}$ are its jumping times, then for all $f \in \mathcal{B}(\mathcal{S})$ and all $n \in \mathbb{N}$, the process:

$$f(X_{t \wedge \tau_n}) - f(X_0) - \int_0^{t \wedge \tau_n} \mathcal{L}[f](X_s) ds$$

is a martingale with respect to the filtration \mathcal{F}^X . This can be proved by induction on n , using the Markov property and the fact that:

$$\mathcal{L}f(X_s) = \sum_{y \neq X_{\tau_n}} \mathcal{L}_{X_{\tau_n}, y} f(y), \quad \forall s \in]\tau_n, \tau_{n+1}[$$

1.3 Strong stationary times and duality

We will now deal more specifically with the theory of strong stationary times, by recalling the main definitions and results of Fill [6] that we will use here. For our purpose, these definitions are given in continuous time, but are easily transposable to the original discrete time case (see e.g. Diaconis and Fill [3]).

Let $(X_t)_{t \geq 0}$ be a Markov chain on a countable set \mathcal{S} , positive recurrent, irreducible and nonexplosive, μ_t its distribution at time $t \geq 0$ and μ_∞ its stationary distribution. A *strong stationary time* T of X is a randomized stopping time relative to X satisfying:

$$\mathcal{L}(X_T | T) = \mu_\infty \quad \text{on the set } \{T < \infty\} \quad (3)$$

(the notation $\mathcal{L}(\cdot)$ stands for the distribution). If in addition $T < +\infty$ almost surely, we say that T is a *finite* strong stationary time.

This tool provides an upper bound for the *separation* between μ_t and μ_∞ , defined by:

$$\mathfrak{s}(t) := \sup_{m \in \mathcal{S}} \left(1 - \frac{\mu_t(m)}{\mu_\infty(m)} \right)$$

since, for every $t \geq 0$:

$$\mathfrak{s}(t) \leq \mathbb{P}(T > t)$$

When the above inequality is an equality, T is called a *time to stationarity*. We recall that the separation itself is an upper bound on total variation distance, so that strong stationary times can be used to study the rate of convergence of processes.

In practice, to construct a strong stationary time, one will often use (and always in this article) the notion of *strong stationary dual*, defined in continuous time by Fill [7], Section 2.1. A process $(X_t^*)_{t \geq 0}$, defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ as X , with countable state space \mathcal{S}^* and absorbing state ∞ , is a strong stationary dual of X if it satisfies, for all $t \geq 0$:

$$\mathcal{L}(X_t^* | \mathcal{F}_\infty^X) = \mathcal{L}(X_t^* | \mathcal{F}_t^X) \quad (4)$$

$$\mathcal{L}(X_t | \mathcal{F}_t^{X^*}) = \mu_\infty \quad \text{on the set } \{X_t^* = \infty\} \quad (5)$$

Relation (4) states that X and X^* are adapted to a same filtration. One can show that the time to absorption of a strong stationary dual is a strong stationary time for X . Conversely, from a strong stationary time, one can construct a strong stationary dual in a canonical way (see Fill [7], Theorem 1). When the time to absorption of the dual is a time to stationarity, the dual is called *sharp*.

If Λ is a transition kernel from \mathcal{S}^* to \mathcal{S} , namely:

$$\forall x^* \in \mathcal{S}^*, \quad \sum_{x \in \mathcal{S}} \Lambda(x^*, x) = \sum_{x \in \mathcal{S}} |\Lambda(x^*, x)| = 1$$

and X^* is a Markov process with values in \mathcal{S}^* , we say that X and X^* are Λ -linked or *intertwined* by Λ if:

$$\mathcal{L}(X_t | \mathcal{F}_t^{X^*}) = \Lambda(X_t^*, \cdot) \quad (6)$$

Fill [7] showed that, in the framework of his *General settings* (that is, X and X^* are nonexplosive, and the sets $\{x \in \mathcal{S} : \Lambda(x^*, x) > 0\}$ are finite for all $x^* \in \mathcal{S}^*$, except maybe the absorbing state ∞), such a coupling always exists as soon as the generators \mathcal{L} and \mathcal{L}^* of X and X^* and their initial distribution μ_0 and μ_0^* respectively are in *algebraic duality*:

$$\begin{cases} \mu_0^* \Lambda = \mu_0 \\ \mathcal{L}^* \Lambda = \Lambda \mathcal{L} \end{cases}$$

In Section 2, we will extend this result to the framework of reversible random walks on countable graphs having finitely many vertices of degree strictly larger than 2.

We consider a discrete graph G , simple, connected and undirected. We call *infinite branches* of G its maximal (for set inclusion) infinite connected subsets whose points have degree 2. A more accurate definition is given in Section 4. We assume then that G is formed by a finite number of infinite branches, plus possibly a finite number of points with finite degrees. If there are only infinite branches, then there is only one and $G = \mathbb{Z}$. This particular case is considered in Section 3. Otherwise, each of the N infinite branches (denoted by Q_i , $i \in I := \llbracket 1, N \rrbracket$) is isomorphic as a graph to \mathbb{N} , and we denote $\varphi_i: \mathbb{N} \rightarrow G$ the corresponding isomorphisms. Such a graph is presented in Figure 1.

Next, assume we are given a generator \mathcal{L} on G , such that for all $x \neq y \in G$, $\mathcal{L}_{x,y} \neq 0$ if and only if x and y are adjacent. Since the graph structure of G will appear only through the generator, we can also choose to consider *a priori* a suitable generator on a countable set, and to endow this latter with a suitable graph structure. That is what will be done in Section 4. We assume that \mathcal{L} is positive recurrent and nonexplosive. This results in the following conditions:

$$\sum_{n \in \mathbb{N}} \mu^i(n) < \infty, \quad \forall i \in \llbracket 1, N \rrbracket \quad (7)$$

$$\sum_{n \in \mathbb{N}} \frac{1}{\mu^i(n) \mathcal{L}_{\varphi_i(n), \varphi_i(n+1)}} \sum_{m=0}^n \mu^i(m) = \infty, \quad \forall i \in \llbracket 1, N \rrbracket \quad (8)$$

where:

$$\mu^i(n) := \prod_{k=0}^{n-1} \frac{\mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}}{\mathcal{L}_{\varphi_i(k+1), \varphi_i(k)}} \quad (9)$$

Under these assumptions, there exists a unique stationary distribution μ , and every process X with generator \mathcal{L} converges weakly to μ . We assume in addition that the diagonal of \mathcal{L} is μ -integrable, and that \mathcal{L} is reversible, so that μ satisfies:

$$\forall p, q \in G, \quad \mu(p) \mathcal{L}_{p,q} = \mu(q) \mathcal{L}_{q,p} \quad (10)$$

In this context, we want to study the strong stationary times of X , and in particular to determine whether there exists a finite one for any initial distribution. To do so, we will construct in a classical way a dual process with values in $\mathcal{P}(G)$ the set of subsets of G , intertwined with X by the kernel Λ defined by:

$$\Lambda(Q, \cdot) = \mu|_Q \quad \forall Q \in \mathcal{P}(G)$$

In the idea of what was done by Diaconis and Fill [3] on finite set or by Miclo [12] or Fill and Lyzinski [5] for diffusions on the real line, this dual process can grow at each step by adding adjacent points. The main problem is that it can also lose any boundary point. To keep it from separating in several parts, when this happens we force it to choose one of its connected components, with a probability proportional to its weight relative to μ . We finally get a connected dual process. By studying its conditions of explosion, and showing that for some initial distributions on X there exists a finite strong stationary time if and only if the one provided by this dual process is finite, we get the main result:

Theorem 1.1. *The process X has a finite strong stationary time, whatever the initial distribution μ_0 , if and only if for every $i \in I$:*

$$\sum_{j=1}^{\infty} \mu^i(j+1) \sum_{k=1}^j \frac{1}{\mu^i(k) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} < \infty \quad (11)$$

Miclo [13] gave some other equivalent condition to this: considering the restriction $\mathcal{L}_{\mathcal{F}}$ of \mathcal{L} on the set:

$$\mathcal{F} := \{f \in \mathbb{L}^2(\mu) : \mu[f] = 0\},$$

then (11) holds for all $i \in I$ if and only if $\mathcal{L}_{\mathcal{F}}^{-1}$ is of trace class.

Remark 1.2. *Unlike what was done for diffusions for instance, we do not prove here that the dual process is sharp for some initial distribution. Actually, we conjecture that there would not exist any dual process that would be Λ -linked and sharp at the same time as soon as the number of infinite branches is greater than two (except of course in the trivial case where the initial distribution is the stationary one).*

2 Results on duality

In the sequel, we will work with the following objects:

- \mathcal{S} and \mathcal{S}^* are two countable, topological spaces (not necessarily endowed with discrete topology). The elements of those sets will generally be denoted by x, y, z for \mathcal{S} and x^*, y^*, z^* for \mathcal{S}^* .
- Λ is a transition kernel from \mathcal{S}^* to \mathcal{S} .
- \mathcal{L} and \mathcal{L}^* are two generators, respectively on \mathcal{S} and \mathcal{S}^* and we assume that \mathcal{L} is nonexplosive.
- $\mathbf{P}(\cdot)$ and $\mathbf{P}^*(\cdot)$ are the minimal solutions of the Kolmogorov equations associated to \mathcal{L} and \mathcal{L}^* respectively.
- μ_0 and μ_0^* are two probability measures on \mathcal{S} and \mathcal{S}^* respectively.
- $(X_t)_{t \geq 0}$ is a Markov process with generator \mathcal{L} and initial distribution μ_0 .

Instead of the hypothesis of finiteness of the sets $\{x : \Lambda(x^*, x) > 0\}$ made in the *General Settings* of Fill [7] (section 2.2B), we will assume that the diagonal of \mathcal{L} is μ -integrable, that is:

$$\sum_{x \in \mathcal{S}} \Lambda(x^*, x) \mathcal{L}_x < +\infty$$

In particular, for all $f \in \mathcal{B}(\mathcal{S})$, the matrix product $\Lambda \mathcal{L} f$ is well defined, associative and finite.

2.1 Minimal processes

Theorem 2.1. *If (μ_0, \mathcal{L}) and (μ_0^*, \mathcal{L}^*) satisfy the double condition:*

$$\begin{cases} \mu_0^* \Lambda = \mu_0 \\ \mathcal{L}^* \Lambda = \Lambda \mathcal{L} \end{cases} \quad (12)$$

then there exists a minimal process $(X_t^)_{t \geq 0}$ on \mathcal{S}^* with generator \mathcal{L}^* and initial distribution μ_0^* , satisfying (4) and (6) until its time of explosion T , that is, for all $t \geq 0$:*

$$\mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}, T > t) = \Lambda(X_t^*, \cdot) \quad (13)$$

$$\mathcal{L}(X_t^* \mid X, T > t) = \mathcal{L}(X_t^* \mid \mathcal{F}_t^X, T > t) \quad (14)$$

For all $f \in \mathcal{B}(\mathcal{S})$ and all $x^* \in \mathcal{S}^*$, we have:

$$\sum_{y^* \in \mathcal{S}^*} \sum_{y \in \mathcal{S}} |\mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(y)| \leq \sup_{y \in \mathcal{S}} f(y) \sum_{y^* \in \mathcal{S}^*} |\mathcal{L}_{x^*, y^*}^*| = 2 \sup_{y \in \mathcal{S}} f(y) \mathcal{L}_{x^*}^*$$

so the matrix product $\mathcal{L}^* \Lambda f$ is always commutative and finite-valued. Under condition (12), we set $\Gamma := \mathcal{L}^* \Lambda = \Lambda \mathcal{L}$ and define a matrix $\bar{\mathcal{L}}$, which will be that of a coupling generator, by:

$$\bar{\mathcal{L}}_{\bar{x}, \bar{y}} := \begin{cases} -\left(\mathcal{L}_x + \mathcal{L}_{x^*}^* + \frac{\Gamma(x^*, x)}{\Lambda(x^*, x)}\right) & , \text{ if } \bar{x} = \bar{y} \\ \mathcal{L}_{x,y} & , \text{ if } y \neq x, y^* = x^* \\ \frac{\mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x)}{\Lambda(x^*, x)} & , \text{ if } y = x, y^* \neq x^* \\ \frac{\mathcal{L}_{x,y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y)}{\Gamma(x^*, y)} & , \text{ if } \Lambda(x^*, y) = 0, \Gamma(x^*, y) > 0 \\ 0 & , \text{ otherwise} \end{cases} \quad (15)$$

with:

$$\bar{x} = (x, x^*), \bar{y} = (y, y^*) \in \bar{\mathcal{S}} := \{(z, z^*) \in \mathcal{S} \times \mathcal{S}^* : \Lambda(z^*, z) > 0\}$$

Some remarks on this definition:

Remark 2.2. (a) $\Lambda(x^*, y) = 0$ implies both $x \neq y$ and $x^* \neq y^*$, so these five cases are mutually exclusive (and obviously exhaustive).

(b) If \mathcal{L}_{x^*, y^*}^* or $\mathcal{L}_{x,y} > 0$ and $\Lambda(x^*, y) = 0$, then with the previous remark one checks that $\Gamma(x^*, y) > 0$. Since the fourth coefficient in (15) is 0 as soon as $\mathcal{L}_{x,y}$ or \mathcal{L}_{x^*, y^*}^* is null, we may replace the fourth case by " $\Lambda(x^*, y) = 0, \mathcal{L}_{x,y} \mathcal{L}_{x^*, y^*}^* > 0$ " without changing the definition of the generator (in that case, we still have $\Gamma(x^*, y) > 0$, and in the cases which have been added to the fifth one, the coefficient $\bar{\mathcal{L}}_{\bar{x}, \bar{y}}$ is 0).

With the previous remarks and a straightforward computation, one easily checks that the coefficients above define a generator on $\bar{\mathcal{S}}$.

Proof of Theorem 2.1. The proof is essentially the same as in Fill [7], Proposition 4 and section 2.3. Most parts are identical and given here for convenience.

Let the condition (12) be fulfilled.

In a first time, we will construct X^* coupled with X , so that the couple (X, X^*) has generator $\bar{\mathcal{L}}$. Then we will check that (X, X^*) satisfies (13) and (14).

Coupling of the two processes: Having observed $X_0 = x_0 \in \mathcal{S}$, we set for each $x_0^* \in \mathcal{S}^*$:

$$X_0^* = x_0^* \quad \text{with probability} \quad \frac{\mu_0^*(x_0^*) \Lambda(x_0^*, x_0)}{\mu_0(x_0)} \quad (16)$$

The distribution of (X_0, X_0^*) is given by:

$$\begin{aligned} \bar{\mu}_0(x_0, x_0^*) &= \mathbb{P}(X_0 = x_0) \mathbb{P}(X_0^* = x_0^* | X_0 = x_0) \\ &= \mu_0(x_0) \frac{\mu_0^*(x_0^*) \Lambda(x_0^*, x_0)}{\mu_0(x_0)} \\ &= \mu_0^*(x_0^*) \Lambda(x_0^*, x_0) \end{aligned} \quad (17)$$

which is a probability measure on $\bar{\mathcal{S}}$ since:

$$\begin{aligned} \sum_{x_0, x_0^*} \bar{\mu}_0(x_0, x_0^*) &= \sum_{x_0^*} \mu_0^*(x_0^*) \sum_{x_0} \Lambda(x_0^*, x_0) \\ &= \sum_{x_0^*} \mu_0^*(x_0^*) \\ &= 1 \end{aligned}$$

In addition, the marginal distribution of X_0^* is μ_0^* .

We set $\tau_0 = 0$, so at this point X^* is constructed up to this time (recall that X is already constructed). We will now construct X^* step by step through times τ_k , $k \in \mathbb{N}$, where τ_k is the k -th transition of the bivariate chain (X, X^*) (in particular, X^* may not jump at time τ_k).

Let $k \geq 1$, and assume by induction that $\tau_0, \dots, \tau_{k-1}$ and $X^*_{[0, \tau_{k-1}]}$ are already constructed. We call σ_k the time of the first transition of X following τ_{k-1} (so it might be equal to σ_{k-1} if X has not jumped at time τ_{k-1}). Let $\bar{x} = (x, x^*) := (X, X^*)_{\tau_{k-1}}$, and let ε_{k-1} be an exponential random variable with parameter $\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x$, independent from $(\epsilon_i)_{i \leq k-2}$ and X . Then:

1. If $\tau_{k-1} + \varepsilon_{k-1} > \sigma_k$, we set $\tau_k = \sigma_k$, and:

$$X^*_{\tau_k} = y^* \quad \text{with probability} \quad \frac{\bar{\mathcal{L}}_{\bar{x}, \bar{y}}}{\bar{\mathcal{L}}_{x, y}} \quad (18)$$

where $y = X_{\tau_k}$ and $\bar{y} = (y, y^*)$.

2. Otherwise, we set $\tau_k = \tau_{k-1} + \varepsilon_{k-1}$, and:

$$X^*_{\tau_k} = y^* \neq x^* \quad \text{with probability} \quad \frac{\bar{\mathcal{L}}_{(x, x^*), (x, y^*)}}{\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x} \quad (19)$$

These probabilities are well defined, since $\mathcal{L}_{x, y}$ and $\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x$ are obviously different from 0 if $X_{\tau_k} = y$ or if $\varepsilon_{k-1} < \infty$ respectively. In addition, in the first case, we get as announced $X^*_{\tau_k} = x^*$ a.s. if $\Lambda(x^*, y) \neq 0$, and:

$$X^*_{\tau_k} = y^* \neq x^* \quad \text{with probability} \quad \frac{\mathcal{L}^*_{x^*, y^*} \Lambda(y^*, y)}{\Gamma(x^*, y)}$$

if $\Lambda(x^*, y) = 0$, the sum of these probabilities being equal to 1.

Thus we get $\tau_k - \tau_{k-1} = \min(\varepsilon_{k-1}, \sigma_k - \tau_{k-1})$. Now, $\varepsilon_{k-1} \sim \exp(\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x)$ and $\sigma_k - \tau_{k-1} \sim \exp(\mathcal{L}_x)$ are independent, so $\tau_k - \tau_{k-1} \sim \exp(\bar{\mathcal{L}}_{\bar{x}})$. In addition, considering the events $A := \{(X, X^*)_{\tau_{k-1}} = \bar{x}\}$ and $B := \{(X, X^*)_{\tau_k} = \bar{y}\}$, we get the following transition probabilities:

- If $y = x$, $y^* \neq x^*$:

$$\begin{aligned} \mathbb{P}(B \mid A) &= \mathbb{P}(X_{\tau_k} = x \mid A) \mathbb{P}(X^*_{\tau_k} = y^* \mid X_{\tau_k} = x, A) \\ &= \mathbb{P}(\varepsilon_{k-1} < \sigma_k - \tau_{k-1} \mid A) \mathbb{P}(X^*_{\tau_k} = y^* \mid \varepsilon_{k-1} < \sigma_k - \tau_{k-1}, A) \\ &= \left(\frac{\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x}{\bar{\mathcal{L}}_{\bar{x}}} \right) \left(\frac{\bar{\mathcal{L}}_{(x, x^*), (x, y^*)}}{\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x} \right) \\ &= \frac{\bar{\mathcal{L}}_{(x, x^*), (x, y^*)}}{\bar{\mathcal{L}}_{\bar{x}}} \end{aligned}$$

- If $y \neq x$:

$$\begin{aligned} \mathbb{P}(B \mid A) &= \mathbb{P}(X_{\tau_k} = y \mid A) \mathbb{P}(X^*_{\tau_k} = y^* \mid X_{\tau_k} = y, A) \\ &= \frac{\mathcal{L}_{x, y}}{\mathcal{L}_x} \mathbb{P}(\varepsilon_{k-1} > \sigma_k - \tau_{k-1} \mid A) \mathbb{P}(X^*_{\tau_k} = y^* \mid X_{\tau_k} = y, A) \\ &= \frac{\mathcal{L}_{x, y}}{\mathcal{L}_x} \left(\frac{\mathcal{L}_x}{\bar{\mathcal{L}}_{\bar{x}}} \right) \left(\frac{\bar{\mathcal{L}}_{\bar{x}, \bar{y}}}{\bar{\mathcal{L}}_{x, y}} \right) \\ &= \frac{\bar{\mathcal{L}}_{\bar{x}, \bar{y}}}{\bar{\mathcal{L}}_{\bar{x}}} \end{aligned}$$

By construction, the process (X, X^*) is a Markov jump process, defined up to its first explosion time, and satisfying (14). We have proven with the previous computations that its generator is indeed $\bar{\mathcal{L}}$.

Check of the intertwining (relation (13)): This will be made in three steps. Firstly, we will show an equation with the generators:

$$\forall x^* \in \mathcal{S}^*, f \in \mathcal{B}(\bar{\mathcal{S}}), \quad \sum_{x \in \mathcal{S}} \sum_{\bar{y} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} f(\bar{y}) = \sum_{\bar{y} \in \bar{\mathcal{S}}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y}) \quad (20)$$

Heuristically, the l.h.s. corresponds to all the transitions from some given $x^* \in \mathcal{S}^*$ to a $x \in \mathcal{S}$, and then from $(x, x^*) \in \bar{\mathcal{S}}$ to $(y, y^*) \in \bar{\mathcal{S}}$, whereas the r.h.s. correspond to a transition from x^* to y^* , then to y , without wondering about x .

In a second time, we will use this equation to prove an inequality involving the corresponding semi-groups:

$$\forall x^* \in \mathcal{S}^*, \bar{y} \in \bar{\mathcal{S}}, \quad \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \bar{\mathbf{P}}_{\bar{x}, \bar{y}}(t) \geq \mathbf{P}_{x^*, y^*}^*(t) \Lambda(y^*, y) \quad (21)$$

and finally, we will deduce by induction on k :

$$\mathbb{P}(X_{t_0}^* = x_0^*, \dots, X_{t_k}^* = x_k^*, X_{t_k} = x_k) = \mu_0^*(x_0^*) \prod_{i=1}^k \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \Lambda(x_k^*, x_k) \quad (22)$$

$\forall k \geq 0, \forall 0 = t_0 \leq \dots \leq t_k, \forall x_0^*, \dots, x_k^* \in \mathcal{S}^*, \forall x_k \in \mathcal{S}$, which implies both (13) and that $\mathbf{P}^*(\cdot)$ is the semi-group of X^* .

Proof of (20): we start by proving the following matrix equality:

$$\sum_{x \in \mathcal{S}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} = \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) \quad (23)$$

Let $x^* \in \mathcal{S}^*, \bar{y} \in \bar{\mathcal{S}}$, with $x^* \neq y^*$. If $\Lambda(x^*, y) \neq 0$, then from the definition of $\bar{\mathcal{L}}$, $\bar{\mathcal{L}}_{\bar{x}, \bar{y}} \neq 0$ implies $y = x$. Hence:

$$\begin{aligned} \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} &= \begin{cases} \Lambda(x^*, y) \bar{\mathcal{L}}_{(y, x^*), \bar{y}} & \text{if } (y, x^*) \in \bar{\mathcal{S}} \\ \sum_{x \neq y} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} & \text{otherwise} \end{cases} \\ &= \begin{cases} \Lambda(x^*, y) \frac{\mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y)}{\Lambda(x^*, y)} & \text{if } (y, x^*) \in \bar{\mathcal{S}} \\ \sum_{x \neq y} \Lambda(x^*, x) \frac{\mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y)}{\Gamma(x^*, y)} & \text{otherwise} \end{cases} \\ &= \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) \end{aligned}$$

And if $x^* \in \mathcal{S}^*, \bar{y} \in \bar{\mathcal{S}}$, with $x^* = y^*$, then:

$$\begin{aligned} \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} &= -\Lambda(x^*, y) \bar{\mathcal{L}}_{(y, x^*), \bar{y}} + \sum_{x \neq y} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} \\ &= -\Lambda(x^*, y) \left(\mathcal{L}_y + \mathcal{L}_{x^*}^* + \frac{\Gamma(x^*, y)}{\Lambda(x^*, y)} \right) + \sum_{x \neq y} \Lambda(x^*, x) \mathcal{L}_{x, y} \\ &= -\Lambda(x^*, y) (\mathcal{L}_y + \mathcal{L}_{x^*}^*) - \Gamma(x^*, y) + \Gamma(x^*, y) + \Lambda(x^*, y) \mathcal{L}_y \\ &= -\mathcal{L}_{x^*}^* \Lambda(x^*, y) \\ &= \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) \end{aligned}$$

Now, for $f \in \mathcal{B}(\bar{\mathcal{S}})$, we have:

$$\begin{aligned}
\sum_{x:\bar{x} \in \bar{\mathcal{S}}} \sum_{\bar{y} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} f(\bar{y}) &= \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \left[\Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{x}} f(\bar{x}) + \sum_{\substack{y:(y, x^*) \in \bar{\mathcal{S}} \\ y \neq x}} \Lambda(x^*, x) \mathcal{L}_{x, y} f((y, x^*)) \right. \\
&\quad \left. + \sum_{\substack{y^*:(x, y^*) \in \bar{\mathcal{S}} \\ y^* \neq x^*}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x) f((x, y^*)) + \sum_{\bar{y}:(y, x^*) \notin \bar{\mathcal{S}}} \frac{\Lambda(x^*, x) \mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})}{\Gamma(x^*, y)} \right] \\
&= \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{x}} f(\bar{x}) + \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \sum_{\substack{y:(y, x^*) \in \bar{\mathcal{S}} \\ y \neq x}} \Lambda(x^*, x) \mathcal{L}_{x, y} f((y, x^*)) \\
&\quad + \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \sum_{\substack{y^*:(x, y^*) \in \bar{\mathcal{S}} \\ y^* \neq x^*}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x) f((x, y^*)) + \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \sum_{\bar{y}:(y, x^*) \notin \bar{\mathcal{S}}} \frac{\Lambda(x^*, x) \mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})}{\Gamma(x^*, y)}
\end{aligned}$$

provided that the four sums in the last equality converge. For the first one, recall that:

$$\bar{\mathcal{L}}_{\bar{x}, \bar{x}} = \mathcal{L}_{x, x} + \mathcal{L}_{x^*, x^*}^* - \frac{\Gamma(x^*, x)}{\Lambda(x^*, x)}$$

so we see that the first sum converges since the diagonal of \mathcal{L} is μ -integrable. The second one converges for the same reason, the third one because f is bounded and:

$$\sum_{y^* \in \mathcal{S}^*} \sum_{x \in \mathcal{S}} |\mathcal{L}_{x^*, y^*}^*| \Lambda(y^*, x) = 2\mathcal{L}_{x^*}^* < +\infty,$$

and the last one because f is bounded and:

$$\begin{aligned}
\sum_{y^* \in \mathcal{S}^*} \sum_{y:(y, x^*) \notin \bar{\mathcal{S}}} \sum_{x \in \mathcal{S}} \frac{\Lambda(x^*, x) |\mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^*| \Lambda(y^*, y)}{\Gamma(x^*, y)} &= \sum_{y^* \in \mathcal{S}^*} \sum_{y:(y, x^*) \notin \bar{\mathcal{S}}} |\mathcal{L}_{x^*, y^*}^*| \Lambda(y^*, y) \\
&\leq 2\mathcal{L}_{x^*}^*
\end{aligned}$$

Getting back to our previous computation, we deduce:

$$\begin{aligned}
\sum_{x:\bar{x} \in \bar{\mathcal{S}}} \sum_{\bar{y} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} f(\bar{y}) &= \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{x}} f(\bar{x}) + \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \sum_{\substack{y:(y, x^*) \in \bar{\mathcal{S}} \\ y \neq x}} \Lambda(x^*, x) \mathcal{L}_{x, y} f((y, x^*)) \\
&\quad + \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \sum_{\substack{y^*:(x, y^*) \in \bar{\mathcal{S}} \\ y^* \neq x^*}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x) f((x, y^*)) + \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \sum_{\bar{y}:(y, x^*) \notin \bar{\mathcal{S}}} \frac{\Lambda(x^*, x) \mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})}{\Gamma(x^*, y)} \\
&= \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{x}} f(\bar{x}) + \sum_{y:(y, x^*) \in \bar{\mathcal{S}}} \sum_{x \neq y \in \mathcal{S}} \Lambda(x^*, x) \mathcal{L}_{x, y} f((y, x^*)) \\
&\quad + \sum_{y^* \in \mathcal{S}^*} \sum_{x:(x, y^*) \in \bar{\mathcal{S}}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x) f((x, y^*)) + \sum_{\bar{y}:(y, x^*) \notin \bar{\mathcal{S}}} \sum_{x \neq y \in \mathcal{S}} \frac{\Lambda(x^*, x) \mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})}{\Gamma(x^*, y)} \\
&= \sum_{\bar{y} \in \bar{\mathcal{S}}} \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} f(\bar{y}) \\
&= \sum_{\bar{y} \in \bar{\mathcal{S}}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})
\end{aligned}$$

by (23).

Proof of (21): Let $\bar{\mathbf{P}}(\cdot)$ be the minimal solution of the Kolmogorov equations associated to $\bar{\mathcal{L}}$. For every $t \geq 0$, we define the matrix $(Q_{x^*, \bar{y}}(t))_{(x^*, \bar{y}) \in \mathcal{S}^* \times \bar{\mathcal{S}}}$ as follows:

$$Q_{x^*, \bar{y}}(t) := \sum_{x:\bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathbf{P}}_{\bar{x}, \bar{y}}(t)$$

Observe that for every $(x^*, \bar{y}) \in \mathcal{S}^* \times \bar{\mathcal{S}}$, the function $Q_{x^*, \bar{y}}(\cdot)$ is the uniform limit of the sequence $(Q_{x^*, \bar{y}}^{(n)}(\cdot))_n$ defined by:

$$Q_{x^*, \bar{y}}^{(n)}(t) := \sum_{x \in A_n} \Lambda(x^*, x) \bar{\mathbf{P}}_{\bar{x}, \bar{y}}(t)$$

where $(A_n)_n$ is a sequence of finite subsets of \mathcal{S} such that $A_n \nearrow \mathcal{S}$. Thus:

$$\begin{aligned} Q'_{x^*, \bar{y}}(t) &= \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathbf{P}}'_{\bar{x}, \bar{y}}(t) \\ &= \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \sum_{\bar{z}} \bar{\mathcal{L}}_{\bar{x}, \bar{z}} \bar{\mathbf{P}}_{\bar{z}, \bar{y}}(t) \\ &= \sum_{\bar{z}} \mathcal{L}_{x^*, z^*}^* \Lambda(z^*, z) \bar{\mathbf{P}}_{\bar{z}, \bar{y}}(t) \\ &= \sum_{z^* \in \mathcal{S}^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(t) \end{aligned}$$

We then get, $\forall (x^*, \bar{y}) \in \mathcal{S}^* \times \bar{\mathcal{S}}$:

$$\begin{aligned} \frac{d}{dt} [\exp(\mathcal{L}_{x^*}^* t) Q_{x^*, \bar{y}}(t)] &= [Q'_{x^*, \bar{y}}(t) + \mathcal{L}_{x^*}^* Q_{x^*, \bar{y}}(t)] \exp(\mathcal{L}_{x^*}^* t) \\ &= \left[\sum_{z^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(t) + \mathcal{L}_{x^*}^* Q_{x^*, \bar{y}}(t) \right] \exp(\mathcal{L}_{x^*}^* t) \\ &= \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(t) \exp(\mathcal{L}_{x^*}^* t) \end{aligned}$$

This sum is finite so the function is continuous in t . We apply the fundamental theorem of calculus, with $Q_{x^*, \bar{y}}(0) = \delta_{x^*, y^*} \Lambda(y^*, y)$:

$$\begin{aligned} \exp(\mathcal{L}_{x^*}^* t) Q_{x^*, \bar{y}}(t) &= \delta_{x^*, y^*} \Lambda(x^*, y) + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(s) \exp(\mathcal{L}_{x^*}^* s) ds \\ Q_{x^*, \bar{y}}(t) &= \delta_{x^*, y^*} \exp(-\mathcal{L}_{x^*}^* t) \Lambda(x^*, y) \\ &\quad + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(s) \exp(\mathcal{L}_{x^*}^* (s - t)) ds \end{aligned} \quad (24)$$

We have expressed Q as a fixed point for the function f defined by:

$$f: \mathcal{M}_{\mathcal{S}^*, \bar{\mathcal{S}}}(\mathcal{C}(\mathbb{R}^+, \mathbb{R}^+)) \rightarrow \mathcal{M}_{\mathcal{S}^*, \bar{\mathcal{S}}}(\mathcal{C}(\mathbb{R}^+, \mathbb{R}^+))$$

$$f(R)_{x^*, \bar{y}}(t) = \delta_{x^*, y^*} \exp(-\mathcal{L}_{x^*}^* t) \Lambda(x^*, y) + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* R_{z^*, \bar{y}}(s) \exp(\mathcal{L}_{x^*}^* (s - t)) ds \quad (25)$$

so that $Q = f(Q)$.

We call:

- Y^* a Markov process with generator \mathcal{L}^*
- J the jumping times chain of the process Y^*
- $P_{x^*, y^*}^{*(N)}(t)$ the probability that Y^* goes from x^* to y^* in time t , and with a maximum of $N - 1$ steps:

$$P_{x^*, y^*}^{*(N)}(t) := \mathbb{P}_{x^*}(Y_t^* = y^*, t < J_N)$$

- $Q_{x^*, \bar{y}}^{*(N)}(t) := P_{x^*, y^*}^{*(N)}(t) \Lambda(y^*, y)$

Then we have:

$$\begin{aligned} P_{x^*, y^*}^{*(1)}(t) &= \mathbb{P}_{x^*}(t < J_1) \delta_{x^*, y^*} \\ &= \exp(-\mathcal{L}_{x^*}^* t) \delta_{x^*, y^*} \end{aligned}$$

and so:

$$Q_{x^*, \bar{y}}^{*(1)}(t) = \delta_{x^*, y^*} \exp(-\mathcal{L}_{x^*}^* t) \Lambda(x^*, y) \leq Q_{x^*, \bar{y}}(t)$$

We will show by induction that $Q^{*(N)} = f(Q^{*(N-1)})$, where f is the function defined in (25), to derive the inequality:

$$Q^* \leq Q \quad (\text{i.e. } Q_{x^*, \bar{y}}^*(t) \leq Q_{x^*, \bar{y}}(t), \forall x^*, \bar{y}, t)$$

where:

$$Q_{x^*, \bar{y}}^*(t) := \mathbf{P}_{x^*, y^*}^*(t) \Lambda(y^*, y)$$

To do so, we condition by $Y_0^* = x^*$ and we separate in two disjoint events:

$$\{Y_t^* = y^*, t < J_N\} = \{Y_t^* = y^*, t < J_1\} \cup \{Y_t^* = y^*, J_1 \leq t < J_N\}$$

the first one having probability $P_{x^*, y^*}^{*(1)}(t)$. For the second one, we condition by J_1 (which is exponentially distributed):

$$\begin{aligned} \mathbb{P}_{x^*}(Y_t^* = y^*, J_1 \leq t < J_N) &= \mathbb{E}_{x^*}(\mathbf{1}_{t < J_N} \mathbf{1}_{Y_t^* = y^*} \mathbf{1}_{J_1 < t}) \\ &= \mathbb{E}_{x^*} \left[\mathbb{E}_{x^*}(\mathbf{1}_{t < J_N} \mathbf{1}_{Y_t^* = y^*} \mid J_1) \mathbf{1}_{J_1 < t} \right] \\ &= \mathbb{E}_{x^*} [\mathbb{P}_{x^*}(t < J_N, Y_t^* = y^* \mid J_1) \mathbf{1}_{J_1 < t}] \end{aligned}$$

then by $Y_{J_1}^*$, which is independent from J_1 and is equal to $z^* \neq x^*$ with probability $\frac{\mathcal{L}_{x^*, z^*}^*}{\mathcal{L}_{x^*}^*}$, and we use the homogeneity of Y^* :

$$\begin{aligned} &\mathbb{P}_{x^*}(Y_t^* = y^*, J_1 < t < J_N) \\ &= \mathbb{E}_{x^*} [\mathbb{P}_{x^*}(t < J_N, Y_t^* = y^* \mid J_1) \mathbf{1}_{J_1 < t}] \\ &= \int_0^t \mathcal{L}_{x^*}^* e^{-s \mathcal{L}_{x^*}^*} \mathbb{P}_{x^*}(t < J_N, Y_t^* = y^* \mid J_1 = s) ds \\ &= \int_0^t \sum_{z^* \in \mathcal{S}^*} \mathcal{L}_{x^*, z^*}^* e^{-s \mathcal{L}_{x^*}^*} \mathbb{P}_{x^*}(t < J_N, Y_t^* = y^* \mid J_1 = s, Y_{J_1}^* = z^*) \mathbb{P}_{x^*}(Y_{J_1}^* = z^*) ds \\ &= \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* e^{-s \mathcal{L}_{x^*}^*} \mathbb{P}_{z^*}(Y_{t-s}^* = y^*, t-s \leq J_{N-1}) ds \\ &= \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* e^{-s \mathcal{L}_{x^*}^*} P_{z^*, y^*}^{*(N-1)}(t-s) ds \\ &= \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* e^{-(t-s) \mathcal{L}_{x^*}^*} P_{z^*, y^*}^{*(N-1)}(s) ds \end{aligned}$$

Regrouping the terms, we get:

$$P_{x^*, y^*}^{*(N)}(t) = \exp(-\mathcal{L}_{x^*}^* t) \delta_{x^*, y^*} + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* e^{-(t-s) \mathcal{L}_{x^*}^*} P_{z^*, y^*}^{*(N-1)}(s) ds$$

so $Q_{x^*, \bar{y}}^{*(N)}(t) = f(Q_{x^*, \bar{y}}^{*(N-1)}(t))$. Since f is non-decreasing, we get by induction $Q_{x^*, \bar{y}}^{*(N)}(t) \leq Q_{x^*, \bar{y}}(t)$. By monotone convergence theorem, $(Q_{x^*, \bar{y}}^{*(N)}(t))_N$ goes to $Q_{x^*, \bar{y}}^*(t) \leq Q_{x^*, \bar{y}}(t)$, which concludes the proof of (21).

When \mathcal{L}^* is nonexplosive, Q^* is a stochastic matrices, the inequality above implies the equality $Q^* = Q$ (since Q is substochastic) and we can show (22) by an easy induction on k (see below). Unfortunately, this argument is no longer valid if \mathcal{L}^* is explosive. The following reasoning shows the nonexplosive case first, then use it to show the explosive case.

Proof of (22): Recall that we want to prove by induction on k :

$$\mathbb{P}(X_{t_0}^* = x_0^*, \dots, X_{t_k}^* = x_k^*, X_{t_k} = x_k) = \mu_0^*(x_0^*) \prod_{i=1}^k \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \Lambda(x_k^*, x_k)$$

T is the explosion time of X^* and actually that of (X, X^*) since we assume that X is nonexplosive. We recall (17):

$$\bar{\mu}_0(x_0, x_0^*) = \mu_0^*(x_0^*) \Lambda(x_0^*, x_0)$$

which gives us immediately the base case for $k = 0$. Suppose now that (22) is true for some $k \geq 0$. Let $x_0^*, \dots, x_{k+1}^* \in \mathcal{S}^*$ and $x_{k+1} \in \mathcal{S}$. Then we have:

$$\begin{aligned} & \mathbb{P}(X_{t_0}^* = x_0^*, \dots, X_{t_{k+1}}^* = x_{k+1}^*, X_{t_{k+1}} = x_{k+1}) \\ &= \sum_{x_k \in \mathcal{S}} \mathbb{P}(X_{t_0}^* = x_0^*, \dots, X_{t_k}^* = x_k^*, X_{t_k} = x_k) \\ & \quad \times \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, X_{t_{k+1}} = x_{k+1} \mid X_{t_0}^* = x_0^*, \dots, X_{t_k}^* = x_k^*, X_{t_k} = x_k) \\ &= \mu_0^*(x_0^*) \prod_{i=1}^k \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \sum_{x_k \in \mathcal{S}} \Lambda(x_k^*, x_k) \bar{\mathbf{P}}_{x_k, x_{k+1}}(t_{k+1} - t_k) \\ &\geq \mu_0^*(x_0^*) \prod_{i=1}^k \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \mathbf{P}_{x_k^*, x_{k+1}^*}^*(t_{k+1} - t_k) \Lambda(x_{k+1}^*, x_{k+1}) \\ &= \mu_0^*(x_0^*) \prod_{i=1}^{k+1} \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \Lambda(x_{k+1}^*, x_{k+1}) \end{aligned}$$

where we have applied successively the law of total probability for the first equality, the Markovianity of (X, X^*) and the induction hypothesis for the second one, and (21) for the inequality.

We have seen that if \mathcal{L}^* is nonexplosive, the inequality is actually an equality and the induction is over. In particular, this shows that $\mathbf{P}^*(\cdot)$ is the semi-group of the marginal X^* and consequently that \mathcal{L}^* is its generator. The process X^* has been built from the trajectories of X , independently from the explosiveness of its generator \mathcal{L}^* . Its jump rates are thus given by the coefficients \mathcal{L}_{x^*, y^*}^* , $x^*, y^* \in \mathcal{S}^*$ and it follows that \mathcal{L}^* is also the generator of the minimal process X^* without the assumption of nonexplosiveness.

Let us assume now that \mathcal{L}^* is explosive. The first and last expressions in the computation above are both positive for any choice of x_i^*, x_{k+1} . Now, their sums in all possible $x_0^*, \dots, x_{k+1}^*, x_{k+1}$ are both equal to $1 - \mathbb{P}(T \leq t_{k+1})$ (since $\mathbf{P}^*(\cdot)$ is the minimal semi-group associated to \mathcal{L}^*), which shows the equality and ends the induction. \square

Corollary 2.3. *If the condition (12) holds, then the process (X, X^*) also satisfies:*

$$\mathcal{L}(X_\tau \mid \mathcal{F}_\tau^{X^*}, \tau < T) = \Lambda(X_\tau^*, \cdot) \quad (26)$$

for every $(\mathcal{F}_t^{X^*})$ -stopping time τ .

Here, the inequality $\tau < T$ has to be understood in $\bar{\mathbb{R}}$. In particular, T and τ might be infinite but the event $\{\tau < T\}$ implies $\{\tau < \infty\}$.

Proof. Let τ be a $(\mathcal{F}_t^{X^*})$ -stopping time, $f: \mathcal{S} \rightarrow \mathbb{R}$ be measurable, continuous and bounded and G be a real-valued random variable, measurable with respect to $\mathcal{F}_\tau^{X^*}$ and bounded. By definition of the conditional distribution, we want to show:

$$\mathbb{E}(G\mathbf{1}_{\tau < T} f(X_\tau)) = \mathbb{E}(G\mathbf{1}_{\tau < T} \Lambda[f](X_\tau^*)) \quad (27)$$

For all $n \in \mathbb{N}$, we set $\tau_n := \frac{1}{n} \lceil n\tau \rceil$, where $\lceil \cdot \rceil$ denote the ceiling function ($\lceil x \rceil$ is the smallest integer greater than or equal to x). We get:

$$\begin{aligned} \{\tau_n \leq t\} &= \{\lceil n\tau \rceil \leq nt\} \\ &= \bigcup_{\substack{m \in \mathbb{N} \\ m \leq nt}} \{m = \lceil n\tau \rceil\} \\ &= \bigcup_{\substack{m \in \mathbb{N} \\ m \leq nt}} \{m - 1 < n\tau \leq m\} \\ &= \bigcup_{\substack{m \in \mathbb{N} \\ \frac{m}{n} \leq t}} \left\{ \frac{m-1}{n} < \tau \leq \frac{m}{n} \right\} \in \mathcal{F}_t^{X^*} \end{aligned}$$

so τ_n is a $(\mathcal{F}_t^{X^*})$ -stopping time. τ_n takes values in $\frac{1}{n}\mathbb{N} \cup \{\infty\}$, and $\tau_n = k/n$ if and only if $\frac{k-1}{n} < \tau \leq \frac{k}{n}$, so:

$$G\mathbf{1}_{\tau_n = k/n} = G\mathbf{1}_{\tau \leq k/n} - G\mathbf{1}_{\tau \leq (k-1)/n}$$

which is measurable with respect to $\mathcal{F}_{k/n}^{X^*}$.

We recall that (13) holds, namely:

$$\forall t > 0, \quad \mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}, t < T) = \Lambda(X_t^*, \cdot)$$

or in other words, by definition, for every function $h: \mathcal{S} \rightarrow \mathbb{R}$ measurable, continuous and bounded, every real-valued random variable H measurable with respect to $\mathcal{F}_t^{X^*}$ and bounded, and every $t > 0$:

$$\mathbb{E}(H\mathbf{1}_{t < T} h(X_t)) = \mathbb{E}(H\mathbf{1}_{t < T} \Lambda[h](X_t^*))$$

Hence:

$$\begin{aligned} \mathbb{E}(\mathbf{1}_{\tau_n < T} Gf(X_{\tau_n})) &= \mathbb{E}\left(\sum_{k=0}^{\infty} \mathbf{1}_{k/n < T} \mathbf{1}_{\tau_n = k/n} Gf(X_{k/n})\right) \\ &= \sum_{k=0}^{\infty} \mathbb{E}\left(\mathbf{1}_{k/n < T} \mathbf{1}_{\tau_n = k/n} Gf(X_{k/n})\right) \\ &= \sum_{k=0}^{\infty} \mathbb{E}\left(\mathbf{1}_{k/n < T} \mathbf{1}_{\tau_n = k/n} G\Lambda[f](X_{k/n}^*)\right) \\ &= \mathbb{E}\left(\sum_{k=0}^{\infty} \mathbf{1}_{k/n < T} \mathbf{1}_{\tau_n = k/n} G\Lambda[f](X_{k/n}^*)\right) \\ &= \mathbb{E}(\mathbf{1}_{\tau_n < T} G\Lambda[f](X_{\tau_n}^*)) \end{aligned}$$

Now, almost surely $\lim_{n \rightarrow \infty} f(X_{\tau_n}) = f(X_\tau)$ by right continuity of the trajectories of (X, X^*) and continuity of f , $\lim_{n \rightarrow \infty} \mathbf{1}_{\tau_n < T} = \mathbf{1}_{\tau < T}$ and $\lim_{n \rightarrow \infty} \Lambda[f](X_{\tau_n}^*) = \Lambda[f](X_\tau^*)$ by continuity of $\Lambda[f]$, so we conclude using dominated convergence:

$$\begin{aligned} \mathbb{E}(G\mathbf{1}_{\tau < T} f(X_\tau)) &= \lim_{n \rightarrow \infty} \mathbb{E}(\mathbf{1}_{\tau_n < T} Gf(X_{\tau_n})) \\ &= \mathbb{E}(G\mathbf{1}_{\tau < T} \Lambda[f](X_\tau^*)) \end{aligned}$$

□

2.2 Duality after explosion

So far, we have focused on minimal processes, without considering what may happen after explosion. This is not enough in general for the Λ -linked process constructed previously to be a strong stationary dual for X .

Indeed, let us consider the simple example where X is a positive recurrent birth-death process on \mathbb{Z} . Proceeding like Diaconis and Fill [3], we want to construct a dual X^* whose states are the intervals of \mathbb{Z} , and linked to X by a kernel Λ which is the stationary distribution restricted to intervals. In that case, the state $[-\infty, +\infty]$ cannot be reached by X^* in finitely many jumps. Since the time when this occurs is the wanted strong stationary time for X , we would like it to be finite, so we need to investigate what happens after explosion. This example will be studied in more details in the next section.

We are now going to specify a little more our framework. We consider a countable topological space E which, for convenience, we will assume to be separated.

Definition 2.4. *We say that a generator $\hat{\mathcal{L}}$ on E is stratified if there exists $N \in \mathbb{N}$ and a N -tuple $(E_i)_{i \in \llbracket 1, N \rrbracket}$ of subsets of E , called stratification of $\hat{\mathcal{L}}$, such that:*

- *The E_i are mutually disjoint.*
- *$\forall i \in \llbracket 1, N \rrbracket$, $\forall x \in E_i$, if $(Y_t^x)_{0 \leq t \leq \tau}$ is a minimal process starting at x with generator $\hat{\mathcal{L}}$, and explosion time τ , then:*

$$\mathbb{P}(Y_t^x \in \bigsqcup_{j=1}^i E_j \mid t < \tau) = 1$$

- *If in addition $\tau < +\infty$, then almost surely there exists $j \leq i$ and ε such that $Y_t^x \in E_j$, $\forall t \in]\tau - \varepsilon, \tau[$, and $\lim_{t \rightarrow \tau} Y_t^x$ exists and is in $\bigsqcup_{k=j+1}^N E_k$. This limit will be denoted by Y_τ^x .*

This definition may be understood as follows: a minimal process whose generator is stratified can jump from one stratum to another only if the rank of the arrival stratum is lower than or equal to that of the departing one. On the other hand, at the explosion time the process must have a limit in a stratum with rank (strictly) greater than that of the stratum where it was right before. Naturally, we say that the process at the explosion time is equal to this limit, and starts again from there. That is done in the following construction.

Remark 2.5. *To satisfy the second condition, we must have $\hat{\mathcal{L}}_{y,z} = 0$ as soon as $y \in E_j$, $z \in E_k$, $k > j$. The Q -matrix $(\hat{\mathcal{L}}_{x,y})_{(x,y) \in E^2}$ is then a lower triangular block matrix, each block corresponding to a different stratum.*

Let $\hat{\mathcal{L}}$ be a stratified generator on E , $(E_i)_{i \in \llbracket 1, N \rrbracket}$ a stratification of $\hat{\mathcal{L}}$, with $N \in \mathbb{N}$, and ν a probability measure on E . We construct a non-minimal process Z with generator $\hat{\mathcal{L}}$ and initial distribution ν as follows:

- We set $T_0 = 0$ and $Z_{T_0} = x$ with probability $\nu(x)$.
- By induction on $i \geq 0$, we assume Z to be constructed up to the time T_i . If $T_i = \infty$, the construction is over. Otherwise, we simulate a minimal process $Y^{(i)}$ with generator $\hat{\mathcal{L}}$ and starting from Z_{T_i} , independently from $Z_{[0, T_i]}$ given Z_{T_i} . Let T'_{i+1} be its explosion time, we set:

$$\begin{aligned} Z_{T_i+t} &= Y_t^{(i)}, \quad \forall t < T'_{i+1} \\ T_{i+1} &= T_i + T'_{i+1} \\ Z_{T_{i+1}} &= Y_{T'_{i+1}}^{(i)} \quad \text{if } T_{i+1} < \infty \end{aligned}$$

Definition 2.6. We call stratified jump process (resp. stratified semi-group) associated to $(\nu, \hat{\mathcal{L}})$ any process Z constructed as above (resp. its transition function). This process is defined up to a double explosion time, possibly finite:

$$T := \lim_{i \rightarrow +\infty} T_i$$

As for minimal processes, we extend it after this double explosion time by setting $X_t = \Delta$ for every $t \geq T$, where Δ is a cemetery point not in E .

By construction, such a process satisfies the strong Markov property, and by Anderson [2] Proposition 1.2.7, its transition function is solution of the backward equation for $\hat{\mathcal{L}}$ (but may not be solution of the forward equation).

Remark 2.7. We could also define the process after the double explosion time by starting again from the limit, up to a triple explosion time, and so on through multiple explosion times, provided that the limits exist. We will see that for our purpose this is not necessary.

At each explosion time, the process goes from one stratum E_i toward another one E_j , with $j > i$, and such a transition is only possible after an explosion.

The next proposition fully justifies the construction we have just made:

Proposition 2.8. Let $\hat{\mathcal{L}}$ be a stratified generator on E and Z a stratified jump process relative to $\hat{\mathcal{L}}$, with explosion times T_n , $n \in \mathbb{N}$. For every bounded, continuous function $f : E \rightarrow \mathbb{R}$, define:

$$M_t^f = f(Z_t) - f(Z_0) - \int_0^t \hat{\mathcal{L}}f(Z_s) ds$$

If $\hat{\mathcal{L}}f$ is bounded then for all $n \in \mathbb{N}$, $(M_{t \wedge T_n}^f)_{t \geq 0}$ is a martingale with respect to the filtration $(\mathcal{F}_t^Z)_{t \geq 0}$. In particular, if the double-explosion time is infinite almost surely, M^f is a martingale.

Proof. We prove by induction on i that the processes $(M_{t \wedge T_i}^f)_{t \geq 0}$ are martingales. For $i = 0$, this is trivial ($T_0 = 0$). Let us assume then that this is true for some $i \geq 0$. We set:

$$Y_t = \begin{cases} Z_{T_i+t} & \text{if } T_i + t < T_{i+1} \\ \Delta & \text{otherwise} \end{cases}$$

Then Y is a minimal process on E with generator $\hat{\mathcal{L}}$. Call τ_n its n -th jumping time, $n \in \mathbb{N}$, we get for all $0 < s < t$:

$$\begin{aligned} \mathbb{E}(M_{T_i+t \wedge \tau_n}^f - M_{T_i+s \wedge \tau_n}^f \mid \mathcal{F}_{T_i+s \wedge \tau_n}^Z) &= \mathbb{E}(M_{T_i+t \wedge \tau_n}^f - M_{T_i+s \wedge \tau_n}^f \mid Z_{T_i+s \wedge \tau_n}) \\ &= \mathbb{E}(f(Y_{t \wedge \tau_n}) - f(Y_{s \wedge \tau_n}) - \int_{s \wedge \tau_n}^{t \wedge \tau_n} \hat{\mathcal{L}}f(Y_r) dr \mid Y_{s \wedge \tau_n}) \\ &= 0 \end{aligned}$$

since Y is a minimal process with generator $\hat{\mathcal{L}}$.

Making n go to infinity, we get by dominated convergence theorem that the process $(M_{(T_i+t) \wedge T_{i+1}}^f)_{t \geq 0}$ is a martingale. Thus:

$$\begin{aligned} \mathbb{E}(M_{t \wedge T_{i+1}}^f - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_s) &= \mathbb{E}(M_{t \wedge T_{i+1}}^f - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\ &= \mathbb{E}(\mathbb{E}(M_{t \wedge T_{i+1}}^f - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_{T_i}) \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\ &= \mathbb{E}(\mathbb{E}(M_{t \wedge T_{i+1}}^f \mid \mathcal{F}_{T_i}) - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\ &= \mathbb{E}(M_{T_i}^f \mathbf{1}_{t > T_i} + M_t^f \mathbf{1}_{t < T_i} - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\ &= \mathbb{E}(M_{t \wedge T_i}^f - M_{s \wedge T_i}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\ &= 0 \end{aligned}$$

If $T_i \xrightarrow{i \rightarrow \infty} \infty$, we use dominated convergence once again to get that $(M_t^f)_{t \geq 0}$ is a martingale. \square

From now on we assume that:

- \mathcal{S} is a discrete topological space.
- There exists a stratification $(\mathcal{S}_i^*)_{i \in \llbracket 1, N \rrbracket}$, $N \in \mathbb{N}$ of \mathcal{L}^* on \mathcal{S}^* . We write $P^*(\cdot)$ for the stratified semi-group associated to \mathcal{L}^* .
- The double-explosion time of any stratified jump process associated to \mathcal{L}^* is infinite almost surely. The matrix $(P_{x^*, y^*}^*(t))_{x^*, y^* \in \mathcal{S}^*}$ is then stochastic for all $t > 0$.
- For all $x \in \mathcal{S}$, $\mathcal{S}^* \ni x^* \rightarrow \Lambda(x^*, x)$ is continuous.

Lemma 2.9. *Under these hypotheses, $(\mathcal{S} \times \mathcal{S}_i^*)_{i \in \llbracket 1, N \rrbracket}$ is a stratification of $\bar{\mathcal{L}}$ on $\mathcal{S} \times \mathcal{S}^*$ endowed with the product topology.*

Proof. This follows trivially from the fact that \mathcal{L} is non-explosive and that the marginals of a minimal process (X, X^*) with generator $\bar{\mathcal{L}}$ have generators \mathcal{L} and \mathcal{L}^* respectively. \square

The next theorem will be an indispensable tool to construct strong stationary duals on countable spaces.

Theorem 2.10. *Under the above hypotheses, if:*

$$\begin{cases} \mu_0^* \Lambda = \mu_0 \\ \mathcal{L}^* \Lambda = \Lambda \mathcal{L} \end{cases}$$

then there exists a stratified jump process $(X_t^)_{t \geq 0}$ on \mathcal{S}^* with generator \mathcal{L}^* and initial distribution μ_0^* , such that for all $t \geq 0$:*

$$\mathcal{L}(X_t | \mathcal{F}_t^{X^*}) = \Lambda(X_t^*, \cdot) \tag{28}$$

$$\mathcal{L}(X_t^* | X) = \mathcal{L}(X_t^* | \mathcal{F}_t^X) \tag{29}$$

Proof. Let X^* be a stratified jump process constructed from minimal processes given by Theorem 2.1. Thus (X, X^*) has generator $\bar{\mathcal{L}}$ (defined in (15)) and initial distribution $\bar{\mu}$ (we recall that $\bar{\mu}(x_0, x_0^*) = \mu_0^*(x_0^*)\Lambda(x_0^*, x_0)$). By construction, X^* satisfies (29). Let us show that it satisfies (28).

We set $T_0 = 0$ and for each $i \geq 0$ we call T_{i+1}^k the k -th jumping time of (X, X^*) following T_i , $k \in \mathbb{N}$ and:

$$T_{i+1} := \lim_{k \rightarrow \infty} T_{i+1}^k$$

the $(i+1)$ -th explosion time. From hypotheses,

$$\lim_{i \rightarrow +\infty} T_i = +\infty$$

We use the convention $T_j^k = +\infty$, $\forall k, \forall j > i$ if $T_i = +\infty$.

First, we prove that the process (X, X^*) is almost surely continuous at each explosion time T_i . Indeed, X^* is continuous at T_i by definition of a stratified process, and X is discontinuous only when it jumps. Let us call σ_n the jumping times of X , assume by induction on n that (X, X^*) is continuous at each explosion time until σ_n , and call T_i the last explosion before σ_n . Let $\tilde{\mathcal{L}}$ be a generator on \mathcal{S}^* defined by $\tilde{\mathcal{L}}_{x^*, y^*} = \bar{\mathcal{L}}_{(X_{\sigma_n}, x^*), (X_{\sigma_n}, y^*)}$ and Y be a stratified process starting from $X_{\sigma_n}^*$ and with generator $\tilde{\mathcal{L}}$. Looking at the construction of X^* we made in the proof of Theorem 2.1, we see that $(X_{\sigma_n+t}^*)_t$ is equal in law to Y until the time σ_{n+1} . Calling τ_k the jumping times of Y before its first explosion, we get:

$$\begin{aligned}
\mathbb{P}(T_{i+1} = \sigma_{n+1} \mid \mathcal{F}_{\sigma_n}^{\bar{X}}) &= \mathbb{P}(\sigma_n + \tau_k \xrightarrow{k \rightarrow \infty} \sigma_{n+1} \mid \mathcal{F}_{\sigma_n}^{\bar{X}}) \\
&= \mathbb{P}(\sigma_n + \tau_1 + \sum_{k=1}^{\infty} (\tau_{k+1} - \tau_k) = \sigma_{n+1} \mid \mathcal{F}_{\sigma_n}^{\bar{X}}) \\
&= 0
\end{aligned}$$

since τ_1 is a continuous random variable, independent from the $\tau_{k+1} - \tau_k$, σ_n and σ_{n+1} conditionally to $\mathcal{F}_{\sigma_n}^{\bar{X}}$. Proceeding the same way (by a second level of induction), starting from T_{i+j} , $j \geq 0$, if it is smaller than σ_{n+1} , we conclude the induction. In conclusion, we get $\sigma_n \neq T_m$, $\forall m, n \geq 0$.

We will now show by induction on i that for all $0 = t_1 < \dots < t_n < \infty$ and all $(x_1^*, \dots, x_n^*, x_n) \in \{\mathcal{S}^*\}^n \times \mathcal{S}$, $n \in \mathbb{N}$, we have:

$$\begin{aligned}
&\mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{T_i}^* = x_n^*, X_{T_i} = x_n \mid t_{n-1} < T_i \leq t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{T_i}^* = x_n^* \mid t_{n-1} < T_i \leq t_n) \Lambda(x_n^*, x_n)
\end{aligned} \tag{30}$$

and:

$$\begin{aligned}
&\mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_{i+1} > t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_{i+1} > t_n) \Lambda(x_n^*, x_n)
\end{aligned} \tag{31}$$

For $i = 0$, we simply consider a minimal process with generator $\bar{\mathcal{L}}$ so the base case is given by the initial distribution of (X, X^*) (for (30)) and Theorem 2.1 (for (31)). Suppose the induction hypothesis true for some $i - 1 \geq 0$. From (31) we deduce:

$$\mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}, T_i > t) = \Lambda(X_t^*, \cdot), \quad \forall t > 0$$

and in the same way as in Corollary 2.3:

$$\mathcal{L}(X_S \mid \mathcal{F}_S^{X^*}) = \Lambda(X_S^*, \cdot)$$

for every stopping time $S < T_i$ almost surely.

We set:

$$\mathcal{A}_i := \bigvee_{k \geq 0} \mathcal{F}_{T_i^k}^{X^*}$$

the σ -algebra generated by the union of $\mathcal{F}_{T_i^k}^{X^*}$.

Let G be a bounded, real-valued random variable, measurable with respect to \mathcal{A}_i , and $f: \mathcal{S} \rightarrow \mathbb{R}$ be a finitely supported function. Then $\Lambda[f]$ is continuous and bounded, and by the convergence of closed martingales theorem, we have $\mathbb{E}(G \mid \mathcal{F}_{T_i^k}^{X^*}) \rightarrow G$ almost surely. Thus using the dominated convergence theorem:

$$\begin{aligned}
\mathbb{E}(Gf(X_{T_i})) &= \lim_{k \rightarrow \infty} \mathbb{E}(\mathbb{E}(G \mid \mathcal{F}_{T_i^k}^{X^*}) f(X_{T_i^k})) \\
&= \lim_{k \rightarrow \infty} \mathbb{E}(\mathbb{E}(G \mid \mathcal{F}_{T_i^k}^{X^*}) \Lambda[f](X_{T_i^k}^*)) \\
&= \mathbb{E}(G \Lambda[f](X_{T_i}^*))
\end{aligned}$$

Now, if we choose $0 = t_1 < \dots < t_n < \infty$ and $(x_1^*, \dots, x_n^*) \in \{\mathcal{S}^*\}^n$, $n \in \mathbb{N}$ then there exists a decreasing sequence $(V_m)_{m \in \mathbb{N}}$ of neighborhoods of x_n^* such that $\bigcap_m V_m = \{x_n^*\}$ and we have:

$$\begin{aligned}
& \{X_{t_1}^* = x_1^*, \dots, X_{T_i}^* = x_n^*, T_i \in]t_{n-1}, t_n]\} \\
&= \{X_{t_1}^* = x_1^*, \dots, \lim_{k \rightarrow \infty} X_{T_i^k}^* = x_n^*, \lim_{k \rightarrow \infty} T_i^k \in]t_{n-1}, t_n]\} \\
&= \bigcap_{m \geq 0} \bigcup_{K \geq 0} \bigcap_{k \geq K} \{X_{t_1}^* = x_1^*, \dots, X_{T_i^k}^* \in V_m, T_i^k \in]t_{n-1}, t_n]\}
\end{aligned}$$

that is:

$$\{X_{t_1}^* = x_1^*, \dots, X_{T_i}^* = x_n^*, T_i \in]t_{n-1}, t_n]\} \in \mathcal{A}_i$$

and we deduce (30). In other words, the intertwining relation is still true "at explosion time T_i ".

By construction, the process $((X, X^*)_{T_i+t})_{t \geq 0}$ has the same law as a minimal process (Y, Y^*) with generator $\bar{\mathcal{L}}$ up to the explosion time T_{i+1} . This process is independent from $((X, X^*)_{T_i \wedge t})_{t \geq 0}$ given $(X, X^*)_{T_i}$ and thanks to what we have done right before, we also know that the initial distribution of this process satisfies:

$$\mathbb{P}(Y_0 = x_0 \mid Y_0^* = x_0^*) = \Lambda(x_0^*, x_0)$$

for every $(x_0, x_0^*) \in \bar{\mathcal{S}}$. From Theorem 2.1 we deduce that for all $0 = t_1 < \dots < t_n < \infty$ and all $(x_1^*, \dots, x_n^*, x_n) \in \{\mathcal{S}^*\}^n \times \mathcal{S}$, $n \in \mathbb{N}$:

$$\begin{aligned}
& \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_{i+1} > t_n, T_i \leq t_1) \\
&= \mathbb{P}(Y_{t_1-T_i}^* = x_1^*, \dots, Y_{t_n-T_i}^* = x_n^*, Y_{t_n-T_i} = x_n \mid T^Y > t_n - T_i, T_i \leq t_1) \\
&= \mathbb{P}(Y_{t_1-T_i}^* = x_1^*, \dots, Y_{t_n-T_i}^* = x_n^* \mid T^Y > t_n - T_i, T_i \leq t_1) \Lambda(x_n^*, x_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_{i+1} > t_n, T_i \leq t_1) \Lambda(x_n^*, x_n)
\end{aligned}$$

where T^Y is the first explosion time of the process (Y, Y^*) . We use it to show that the intertwining relation is still valid after the explosion time T_i . Indeed, if $k \in \llbracket 1, n-1 \rrbracket$, then using:

- i. the Markovianity of (X, X^*)
- ii. the independence of X_{T_i} and $X_{[0, T_i]}^*$ given $X_{T_i}^*$
- iii. the previous computation

we get:

$$\begin{aligned}
& \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&= \sum_{\bar{x} \in \bar{\mathcal{S}}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n, \bar{X}_{T_i} = \bar{x} \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\stackrel{i.}{=} \sum_{\bar{x} \in \bar{\mathcal{S}}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, \bar{X}_{T_i} = \bar{x} \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid \bar{X}_{T_i} = \bar{x}, T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\stackrel{ii.}{=} \sum_{\bar{x} \in \bar{\mathcal{S}}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{T_i} = x \mid X_{T_i}^* = x^*, T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \frac{\mathbb{P}(X_{T_i} = x, X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}])}{\mathbb{P}(X_{T_i} = x \mid X_{T_i}^* = x^*, T_i \in]t_k, t_{k+1}], T_{i+1} > t_n)} \\
&= \sum_{\bar{x} \in \bar{\mathcal{S}}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{T_i} = x, X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&= \sum_{x^* \in \mathcal{S}^*} \sum_{x \in \mathcal{S}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{T_i} = x, X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&= \sum_{x^* \in \mathcal{S}^*} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\stackrel{iii.}{=} \sum_{x^* \in \mathcal{S}^*} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n}^* = x_n^* \mid X_{T_i}^* = x^*, T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \Lambda(x_n^*, x_n)
\end{aligned}$$

Keeping $\Lambda(x_n^*, x_n)$ apart, we do the same steps backward to get that this quantity is equal to:

$$\mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \Lambda(x_n^*, x_n)$$

so the only thing left is to sum on all possible intervals for T_i :

$$\begin{aligned}
& \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_{i+1} > t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_{i+1} > t_n, T_i > t_n) \mathbb{P}(T_i > t_n \mid T_{i+1} > t_n) \\
&\quad + \sum_{k=1}^{n-1} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n, t_k < T_i \leq t_{k+1} \mid T_{i+1} > t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_i > t_n) \Lambda(x_n^*, x_n) \mathbb{P}(T_i > t_n \mid T_{i+1} > t_n) \\
&\quad + \sum_{k=1}^{n-1} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, t_k < T_i \leq t_{k+1} \mid T_{i+1} > t_n) \Lambda(x_n^*, x_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_{i+1} > t_n) \Lambda(x_n^*, x_n)
\end{aligned}$$

which concludes the induction. We deduce:

$$\mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}, T_i > t) = \Lambda(X_t^*, \cdot)$$

for every $i \in \mathbb{N}$, that is:

$$\mathbb{E}(\mathbf{1}_{T_i > t} G f(X_t)) = \mathbb{E}(\mathbf{1}_{T_i > t} G \Lambda[f](X_t^*))$$

for every bounded continuous function f and every bounded random variable G , measurable with respect to $\mathcal{F}_t^{X^*}$. Now, $\lim_{i \rightarrow \infty} \mathbf{1}_{T_i > t} = 1$ almost surely and we conclude by dominated convergence:

$$\mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}) = \Lambda(X_t^*, \cdot)$$

□

Corollary 2.11. *The relation (28) may be replaced by:*

$$\mathcal{L}(X_T \mid \mathcal{F}_T^{X^*}, T < \infty) = \Lambda(X_T^*, \cdot) \quad (32)$$

for every $\mathcal{F}_t^{X^*}$ -stopping time T .

Proof. Condition (32) trivially implies (28). The converse is the same as for Corollary 2.3. □

3 Application to birth-death processes

Let $(X_t)_{t \geq 0}$ be a birth-death process on \mathbb{Z} , with death rates $(a_n)_{n \in \mathbb{Z}}$ and birth rates $(b_n)_{n \in \mathbb{Z}}$, all positive. We assume that the coefficients are such that X is nonexplosive and positive recurrent, that is (cf Anderson [2] Chapter 8):

$$\sum_{i=-\infty}^{+\infty} \mu(i) < \infty \quad (33)$$

$$\begin{cases} \sum_{i=0}^{+\infty} \frac{1}{b_i \mu(i)} = \infty \\ \sum_{i=-\infty}^0 \frac{1}{b_i \mu(i)} = \infty \end{cases} \quad (34)$$

where μ is the stationary distribution, uniquely defined by the relations:

$$\mu(n)b_n = \mu(n+1)a_{n+1} \quad (35)$$

$$\sum_{n \in \mathbb{Z}} \mu(n) = 1 \quad (36)$$

namely, for every $n \in \mathbb{N}^*$:

$$\mu(n) = \mu(0) \prod_{i=0}^{n-1} \frac{b_i}{a_{i+1}}, \quad \mu(-n) = \mu(0) \prod_{i=0}^{n-1} \frac{a_{-i}}{b_{-i-1}}$$

$$\text{with } \mu(0) = \frac{1}{1 + \sum_{n \in \mathbb{N}^*} \left(\prod_{i=0}^{n-1} \frac{b_i}{a_{i+1}} + \prod_{i=0}^{n-1} \frac{a_{-i}}{b_{-i-1}} \right)}.$$

We assume that the coefficients are such that:

$$\sum_{n \in \mathbb{Z}} \mu(n)b_n < \infty$$

The generator $(\mathcal{L}_{i,j})_{(i,j) \in \mathbb{Z}^2}$ of this process X is defined by:

$$\forall i, j \in \mathbb{Z}, \quad \mathcal{L}_{i,j} = \begin{cases} b_i & \text{if } j = i + 1 \\ a_i & \text{if } j = i - 1 \\ -a_i - b_i & \text{if } j = i \\ 0 & \text{otherwise} \end{cases} \quad (37)$$

We denote μ_t the distribution of X_t , $\forall t \geq 0$.

Starting with this process, we want to construct a strong stationary dual of X , with values in:

$$\mathbf{E} := \{(p, q) \in (\mathbb{Z} \cup \{-\infty\}) \times (\mathbb{Z} \cup \{+\infty\}) : p \leq q\}$$

the set of all interval of \mathbb{Z} , linked to X by the kernel Λ defined by:

$$\Lambda((p, q), n) = \delta_n(\llbracket p, q \rrbracket) \frac{\mu(n)}{\mu(\llbracket p, q \rrbracket)}, \quad \forall (p, q) \in \mathbf{E}, \quad \forall n \in \mathbb{Z}$$

Here and in the sequel, for $(p, q) \in \mathbf{E}$, $n \in \llbracket p, q \rrbracket$ means $n \in [p, q] \cap \mathbb{Z}$ (we exclude $n = \pm\infty$).

We endow \mathbf{E} with the topology generated by the following open sets:

$$\begin{aligned} & \{(p, q)\} \\ & \{(p, q + i), i \in \mathbb{N}\} \cup \{(p, +\infty)\} \\ & \{(p - i, q), i \in \mathbb{N}\} \cup \{(-\infty, q)\} \\ & \{(-\infty, q + i), i \in \mathbb{N}\} \cup \{\mathbb{Z}\} \\ & \{(p - i, +\infty), i \in \mathbb{N}\} \cup \{\mathbb{Z}\} \\ & \{(p - i, q + j), i, j \in \mathbb{N}\} \cup \{\mathbb{Z}\} \end{aligned}$$

with $-\infty < p \leq q < +\infty$. We set:

$$\mathbf{E}_1 := \{(p, q) \in \mathbb{Z}^2 : p \leq q\}$$

$$\mathbf{E}_2 := \{(-\infty, q) \in \mathbf{E}\} \cup \{(p, +\infty) \in \mathbf{E}\}$$

$$\mathbf{E}_3 := \{\mathbb{Z}\}$$

Let \mathcal{L}^* be a generator on \mathbf{E} defined as follows:

$$\mathcal{L}^*((p, q), (p_i, q_i)) = \begin{cases} \frac{\mu(\llbracket p-1, q \rrbracket)}{\mu(\llbracket p, q \rrbracket)} b_{p-1} =: a'_{p,q} & \text{if } (p_i, q_i) = (p-1, q) \\ \frac{\mu(\llbracket p, q+1 \rrbracket)}{\mu(\llbracket p, q \rrbracket)} a_{q+1} =: b_{p,q} & \text{if } (p_i, q_i) = (p, q+1) \\ \frac{\mu(\llbracket p+1, q \rrbracket)}{\mu(\llbracket p, q \rrbracket)} a_p =: b'_{p,q} & \text{if } (p_i, q_i) = (p+1, q) \\ \frac{\mu(\llbracket p, q-1 \rrbracket)}{\mu(\llbracket p, q \rrbracket)} b_q =: a_{b,q} & \text{if } (p_i, q_i) = (p, q-1) \\ 0 & \text{for all other } (p_i, q_i) \neq (p, q) \end{cases} \quad (38)$$

for all $(p, q) \in \mathbf{E}$, with the convention $a_\infty = b_\infty = a_{-\infty} = b_{-\infty} = 0$. Then we get:

Proposition 3.1. *There exists a measure μ_0^* on \mathbf{E} and a minimal process X^* with initial distribution μ_0^* and generator \mathcal{L}^* such that X^* is Λ -linked to X .*

Proof. From Theorem 2.1, it is enough to have $\mu_0^* \Lambda = \mu_0$ and $\mathcal{L}^* \Lambda = \Lambda \mathcal{L}$. For the first one it is immediate if we choose the measure whose support is the set of singletons: $\mu_0^*((m, n)) = \delta_{m,n} \mu_0(n)$. For the second one: $\forall (p, q) \in \mathbf{E}, \forall n \in \mathbb{Z}$

$$\begin{aligned} (\Lambda \mathcal{L})_{(p,q),n} &= \sum_{m \in \mathbb{Z}} \Lambda((p, q), m) \mathcal{L}(m, n) \\ &= \sum_{m \in \llbracket p, q \rrbracket} \Lambda((p, q), m) \mathcal{L}(m, n) \\ &= \frac{\mu(n-1)}{\mu(\llbracket p, q \rrbracket)} \delta_{n-1}(\llbracket p, q \rrbracket) b_{n-1} \\ &\quad + \frac{\mu(n+1)}{\mu(\llbracket p, q \rrbracket)} \delta_{n+1}(\llbracket p, q \rrbracket) a_{n+1} - \frac{\mu(n)}{\mu(\llbracket p, q \rrbracket)} \delta_n(\llbracket p, q \rrbracket) (b_n + a_n) \\ &= \frac{\mu(n)[b_n(\delta_{n,p-1} - \delta_{n,q}) + a_n(\delta_{n,q+1} - \delta_{n,p})]}{\mu(\llbracket p, q \rrbracket)} \end{aligned}$$

and:

$$\begin{aligned}
(\mathcal{L}^* \Lambda)_{(p,q),n} &= \sum_{(p',q') \in \mathbf{E}} \mathcal{L}^*((p,q), (p',q')) \Lambda((p',q'), n) \\
&= \frac{\mu(n)}{\mu(\llbracket p, q \rrbracket)} (b_{p-1} \delta_n(\llbracket p-1, q \rrbracket) - \delta_n(\llbracket p, q \rrbracket) b_{p-1} \\
&\quad + a_{q+1} \delta_n(\llbracket p, q+1 \rrbracket) - \delta_n(\llbracket p, q \rrbracket) a_{q+1} \\
&\quad + a_p \delta_n(\llbracket p+1, q \rrbracket) - \delta_n(\llbracket p, q \rrbracket) a_p \\
&\quad + b_q \delta_n(\llbracket p, q-1 \rrbracket) - \delta_n(\llbracket p, q \rrbracket) b_q) \\
&= \frac{\mu(n) [b_n (\delta_{n,p-1} - \delta_{n,q}) + a_n (\delta_{n,q+1} - \delta_{n,p})]}{\mu(\llbracket p, q \rrbracket)} \\
&= (\Lambda \mathcal{L})_{(p,q),n}
\end{aligned}$$

□

So we have a minimal process $(X_t^*) = (P_t, Q_t) \in \mathbf{E}$. Its bounds P and Q evolve as birth-death processes individually not Markovian on \mathbb{Z} , until a possible finite explosion time T . Here, the only absorbing state for X^* is \mathbb{Z} , but it is not reachable in finitely many jumps (if X^* starts from another state). We hence have to consider a stratified jump process associated to \mathcal{L}^* , which is made possible by:

Proposition 3.2. \mathcal{L}^* is stratified on \mathbf{E} , and $(\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3)$ is a stratification of \mathcal{L}^* .

Proof. For a minimal process $(X_t^*)_t = (P_t, Q_t)_t$ with generator \mathcal{L}^* and initial distribution μ_0^* , we call T its first explosion time and we set:

$$\begin{aligned}
T_1 &:= \begin{cases} T & \text{if } T \text{ is an explosion time for } P \\ +\infty & \text{otherwise} \end{cases} \\
T_2 &:= \begin{cases} T & \text{if } T \text{ is an explosion time for } Q \\ +\infty & \text{otherwise} \end{cases}
\end{aligned}$$

so that $T = T_1 \wedge T_2$.

We will show that $\lim_{t \rightarrow T} Q_t = +\infty$ a.s. if $T = T_2$, and we deduce by symmetry that $\lim_{t \rightarrow T} P_t = -\infty$ if $T = T_1$. To do so, we consider the process $X^* = (P, Q)$ coupled with a (homogeneous) birth-death process Z with same initial distribution as Q , corresponding to the evolution of the right bound of X^* when the left bound is in $-\infty$. Such a coupling is given by:

Lemma 3.3. If $T = T_2$, there exists a coupling (X^*, Z) such that $Q_t \geq Z_t$, $\forall 0 \leq t \leq T$.

Proof. We are going to construct $(Z_t)_{t \geq 0}$, such that $(-\infty, Z)$ is a Markov process with generator \mathcal{L}^* . We simulate $X_0^* = (P_0, Q_0)$ with its initial distribution μ_0^* , and we set $Z_0 = Q_0$. The process Z is a birth-death process with death and birth rates respectively:

$$\tilde{b}_n = \frac{\mu(\llbracket -\infty, n+1 \rrbracket)}{\mu(\llbracket -\infty, n \rrbracket)} a_{n+1} \quad \text{and} \quad \tilde{a}_n = \frac{\mu(\llbracket -\infty, n-1 \rrbracket)}{\mu(\llbracket -\infty, n \rrbracket)} b_n$$

An elementary computation shows that for all $(p, q) \in \mathbf{E}$, $b_{p,q} \geq \tilde{b}_q$ and $a_{p,q} \leq \tilde{a}_q$ that is, the births are on average quicker for the process $(Q_t)_t$, and the deaths are slower. To make the coupling, we proceed by letting the two processes evolve separately and independently as long as $Q_t > Z_t$. Then, when they join, we "force" Z_t to go down if Q_t does, and Q_t to go up if Z_t does.

For this purpose, we will define by iteration the sequence $(T_i)_i$ of junction times of Q_t and Z_t , and $(T'_i)_i$ that of the jumping times following the junction (if $Q_{T'_i} = Z_{T'_i}$ we set $T_{i+1} = T'_i$), and we will describe the behavior of the coupling (X^*, Z) between those times. Let $T_0 = 0$. Assume that (X^*, Z) is constructed until the time T_i , for some $i \geq 0$. We define T'_i as a random variable such

that $T'_i - T_i$ is exponentially distributed with parameter $\lambda_i := \tilde{a}_{Q_{T_i}} + b_{P_{T_i}, Q_{T_i}} + a'_{P_{T_i}, Q_{T_i}} + b'_{P_{T_i}, Q_{T_i}}$. We impose $Q_t = Z_t = Q_{T_i}$ and $P_t = P_{T_i}$, $\forall T_i < t < T'_i$, then:

$$((P_{T'_i}, Q_{T'_i}), Z_{T'_i}) = \begin{cases} ((P_{T_i}, Q_{T_i} - 1), Z_{T_i} - 1) & \text{with probability } \frac{a_{P_{T_i}, Q_{T_i}}}{\lambda_i} \\ ((P_{T_i}, Q_{T_i} + 1), Z_{T_i} + 1) & \text{with probability } \frac{\tilde{b}_{Q_{T_i}}}{\lambda_i} \\ ((P_{T_i}, Q_{T_i}), Z_{T_i} - 1) & \text{with probability } \frac{\tilde{a}_{Q_{T_i}} - a_{P_{T_i}, Q_{T_i}}}{\lambda_i} \\ ((P_{T_i}, Q_{T_i} + 1), Z_{T_i}) & \text{with probability } \frac{b_{P_{T_i}, Q_{T_i}} - \tilde{b}_{Q_{T_i}}}{\lambda_i} \\ ((P_{T_i} - 1, Q_{T_i}), Z_{T_i}) & \text{with probability } \frac{a'_{P_{T_i}, Q_{T_i}}}{\lambda_i} \\ ((P_{T_i} + 1, Q_{T_i}), Z_{T_i}) & \text{with probability } \frac{b'_{P_{T_i}, Q_{T_i}}}{\lambda_i} \end{cases}$$

Next, X_t^* and Z_t evolve independently between times T'_i and $T_{i+1} := \inf\{t \geq T'_i : Q_t = Z_t\}$. We clearly have $Q_t \geq Z_t$, $\forall 0 \leq t \leq T$. Moreover, one checks that the marginals are the ones we want: write the coupling generator from the previous coefficients, and apply it to functions depending only on one variable. \square

It only remains to check:

$$Z_t \xrightarrow{t \rightarrow +\infty} +\infty \quad \text{a.s.} \quad (39)$$

$(-\infty, Z)$ is equal in law to a process \tilde{X}^* , Λ -linked to some \tilde{X} with generator \mathcal{L} . The latter is positive recurrent, and for all $t \geq 0$, $\tilde{X}_t \in \tilde{X}_t^*$ a.s., so $Z_t > 0$ infinitely often (considering the underlying jump chain). This implies in particular that $Z_t \xrightarrow{t \rightarrow +\infty} -\infty$ is impossible a.s..

If we consider the same process \tilde{Z}_t , forced to stay on \mathbb{N} by setting $a_0 = 0$, then the condition for $\tilde{Z}_t \xrightarrow{t \rightarrow +\infty} +\infty$ is:

$$\sum_{i=0}^{\infty} \prod_{j=1}^i \frac{\tilde{a}_j}{\tilde{b}_j} < \infty$$

Now:

$$\begin{aligned} \sum_{i=0}^{\infty} \prod_{j=1}^i \frac{\tilde{a}_j}{\tilde{b}_j} &= \sum_{i=0}^{\infty} \prod_{j=1}^i \frac{\mu(\llbracket -\infty, j-1 \rrbracket) b_j}{\mu(\llbracket -\infty, j+1 \rrbracket) a_{j+1}} \\ &\leq \sum_{i=0}^{\infty} \prod_{j=1}^i \frac{b_j}{a_{j+1}} < \infty \end{aligned}$$

Consequently, $\tilde{Z}_t \xrightarrow{t \rightarrow +\infty} \infty$ almost surely, and so does Z_t . Indeed, as long as $Z_t > 0$, the process Z behaves like \tilde{Z} . If we set $\tau_0 = 0$, $\sigma_0 = \inf\{t \geq \tau_0, Z_t = 0\}$, and for every $i > 0$:

$$\tau_i = \inf\{t > \sigma_{i-1}, Z_t = 1\} \quad \text{and} \quad \sigma_i = \inf\{t > \tau_i, Z_t = 0\}$$

(with the convention $\inf \emptyset = +\infty$), we have $\mathbb{P}(\tau_i < \infty \mid \sigma_{i-1} < \infty) = 1$ since $Z_t > 0$ infinitely often and:

$$\mathbb{P}(\sigma_i = \infty \mid \tau_i < \infty) = \mathbb{P}(\tilde{Z}_t \neq 0, \forall t > 0 \mid \tilde{Z}_0 = 1) > 0$$

So there exists almost surely a time τ_i for which $Z_t > 0$ for all $t \geq \tau_i$ (succession of independent Bernoulli variables). Since:

$$\mathcal{L}(Z_{\tau_i+} \mid \tau_i < \infty, \sigma_i = \infty) = \mathcal{L}(\tilde{Z} \mid \tilde{Z}_0 = 1, \tilde{Z}_t > 0, \forall t > 0)$$

it follows that $Z_t \xrightarrow{t \rightarrow +\infty} \infty$ almost surely.

If on the opposite $T < T_2$, then Q does not explode at time T and thus admits a last jumping time T_k before T . The same holds for P , and we conclude that $\lim_{t \rightarrow T} X_t^*$ exists almost surely.

One easily checks that all three conditions for $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ to be a stratification of \mathcal{L}^* are satisfied. \square

By Theorem 2.10, we can now consider a stratified jump process X^* with generator \mathcal{L}^* and initial distribution μ_0^* satisfying $\mu_0^* \Lambda = \mu_0$, such that X^* is Λ -linked to X . The interest of this process is that the existence of a finite strong stationary time for an initial distribution μ_0 of X is equivalent to the absorption in finite time of X^* in \mathbb{Z} , for at least one μ_0^* . The first implication derives directly from the definition of Λ and Corollary 2.11, the converse comes from the following proposition:

Proposition 3.4. *If $\mu_0 = \mu_{\llbracket -\infty, 0 \rrbracket}$ and $X_0^* = (-\infty, 0)$ a.s., then the dual X^* is sharp.*

For the proof, refer to Miclo [12], Lemma 26.

So, in that case, the time to absorption in \mathbb{Z} of X^* is a time to stationarity and its almost sure finiteness is equivalent to the existence of a finite strong stationary time (at least for that initial distribution μ_0).

We are finally able to state the main result of this section:

Theorem 3.5. *There exists a finite strong stationary time of X , whatever the initial distribution, if and only if:*

$$\sum_{i=1}^{\infty} \mu(i+1) \sum_{j=1}^i \frac{1}{\mu(j)b_j} < \infty \quad (40)$$

$$\sum_{i=1}^{\infty} \mu(-i-1) \sum_{j=1}^i \frac{1}{\mu(-j)a_{-j}} < \infty \quad (41)$$

Proof. • Necessary condition:

We consider the same assumption as in Proposition 3.4, that is $\mu_0 = \mu_{\llbracket -\infty, 0 \rrbracket}$ and $X_0^* = \llbracket -\infty, 0 \rrbracket$ a.s.. We saw that in this case, a dual X^* with generator \mathcal{L}^* is sharp, so if there exists a strong stationary time then X^* is absorbed in finite time almost surely. But this happens if and only if the series

$$\sum_{i=1}^{\infty} \prod_{j=1}^i \frac{\tilde{a}_j}{\tilde{b}_j} \sum_{k=1}^i \prod_{l=1}^k \frac{\tilde{b}_{l-1}}{\tilde{a}_l}$$

converges. Indeed, this is the criterion for the explosion in finite time of a birth-death process on \mathbb{N} (see e.g. Anderson [2], Section 8.1), that we extend to a process on \mathbb{Z} with values in \mathbb{N} infinitely often, in the same way as for (39). Using the fact that $\mu(\llbracket -\infty, j \rrbracket)$ is uniformly bounded on $j \in \mathbb{N}$, this condition can be restated as follows:

$$\begin{aligned} & \sum_{i=1}^{\infty} \prod_{j=1}^i \frac{b_j}{a_{j+1}} \sum_{k=1}^i \prod_{l=1}^k \frac{a_l}{b_l} < \infty \\ \Leftrightarrow & \sum_{i=1}^{\infty} \mu(i+1) \sum_{j=1}^i \frac{1}{\mu(j)b_j} < \infty \end{aligned}$$

This proves that (40) is a necessary condition. We do the same for (41), taking $\mu_{\llbracket 0, +\infty \rrbracket}$ as an initial distribution for X , $X_0^* = \llbracket 0, +\infty \rrbracket$ almost surely and considering the left bound of the dual process X^* .

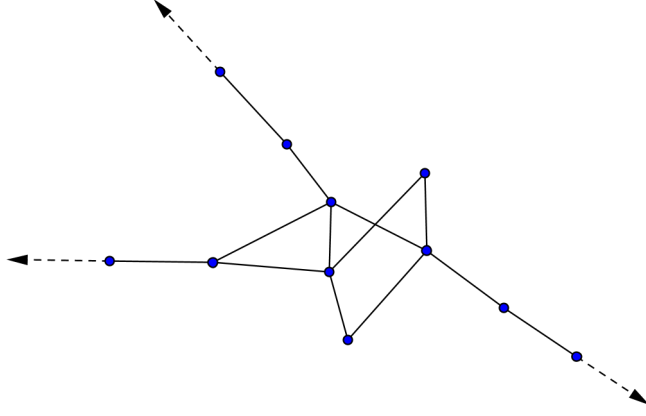


Figure 1: Example of considered graph

- **Sufficient condition:**

Conversely, let $\mu_0^* := \sum_{p \in \mathbb{Z}} \mu_0(p) \delta_{\{p\}}$. We have $\mu_0 = \mu_0^* \Lambda$ so there exists a Λ -linked dual X^* of X with generator \mathcal{L}^* and initial distribution μ_0^* . By Lemma 3.3, we can couple it with a process $(-\infty, Z)$ starting from $\llbracket -\infty, 0 \rrbracket$ with generator \mathcal{L}^* , so that $(-\infty, Z) \subset X^*$. Thus when Z reaches $+\infty$, X^* is absorbed. This happens in finite time if (40) holds. We proceed similarly for (41). □

4 Random walk on a graph

The previous case may be seen as a random walk on a graph, where \mathbb{Z} is a graph made of two infinite branches connected in 0. We want to extend the result of the previous section to graphs having finitely many infinite branches, and another finite number of vertices. Such a graph is represented in Figure 1.

Thus we consider a generator \mathcal{L} on a countable set G , irreducible and such that for all $x \in G$, the set $\{y \in G : \mathcal{L}_{x,y} > 0 \text{ or } \mathcal{L}_{y,x} > 0\}$ is finite. Let $(X_t)_{t \geq 0}$ be a minimal process with generator \mathcal{L} and call μ_0 its initial distribution. We endow G with the structure of a simple, undirected, weighted and connected graph, for which two distinct vertices p and $q \in G$ are adjacent if and only if $\mathcal{L}_{p,q} \neq 0$ or $\mathcal{L}_{q,p} \neq 0$ (each vertex then has finite degree).

We assume that \mathcal{L} is nonexplosive, positive recurrent and reversible, so that there exists a unique stationary distribution μ , and we assume in addition that the sum $\sum \mu(p) \mathcal{L}_p$ is finite. We recall that μ satisfies (10):

$$\mu(q) \mathcal{L}_{q,p} = \mu(p) \mathcal{L}_{p,q}, \quad \forall p, q \in G$$

and in particular, $\mathcal{L}_{p,q} \neq 0$ implies $\mathcal{L}_{q,p} \neq 0$. This equality as much as $\sum_q \mathcal{L}_{p,q} = 0$ will be the basis of all computations.

Let us start by setting the vocabulary and notations that will be used in the sequel.

Definition 4.1. (i) For every $Q \subset G$, we will write Q^c for the complementary set of Q in G .

(ii) Two vertices of G are adjacent if they are connected by an edge.

For $Q \subset G$ and $q \in Q$, we denote by $V_Q(q)$ the set of all points of Q adjacent to q , $V(q) := V_G(q)$ and $V(Q) := \{q' \in V_{Q^c}(q), q \in Q\}$.

(iii) The degree of a vertex is the number of its adjacent points.

(iv) A path between two vertices $p, q \in G$ is a sequence of vertices $p = p_1, \dots, p_n = q$ such that p_i and p_{i+1} are adjacent for every $1 \leq i < n$, and that all p_i are distinct.

(v) A subset Q of G is connected if between any two vertices of Q there exists a path in Q . It is convex if any path between two vertices of Q is in Q .

- (vi) If $Q' \subset Q \subset G$, Q' is a connected component of Q if Q' is connected and maximal (for set inclusion) among connected subsets of Q . We write $\mathcal{C}(Q)$ for the set of all connected components of Q .
- (vii) We call center of G and denote by C_G the union of all paths between any two vertices with degrees different from two (namely, the convex envelope of these vertices).
- (viii) We call infinite branch any connected component (necessarily infinite) of $G \setminus C_G$. We denote Q_i , $i \in I$ these infinite branches of G .
- (ix) To each infinite branch Q_i , we add an infinite point, written Δ_i . We write $\bar{Q}_i = Q_i \cup \{\Delta_i\}$.
- (x) In the case where C_G is non empty, if we endow \mathbb{N} with a structure of graph for which two points n and $m \in \mathbb{N}$ are adjacent if and only if $|n - m| = 1$, then each Q_i is isomorphic as a graph to \mathbb{N} and we call $\varphi_i: \mathbb{N} \cup \{\infty\} \rightarrow \bar{Q}_i$ the unique corresponding isomorphism, completed by setting $\varphi_i(\infty) = \Delta_i$. We will also use the following notation:

$$\varphi_i(\llbracket p, \infty \rrbracket) := \bigcup_{q \geq p} \{\varphi_i(q)\} \cup \{\Delta_i\}$$

- (xi) We write as in (9):

$$\mu^i(n) := \prod_{k=0}^{n-1} \frac{\mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}}{\mathcal{L}_{\varphi_i(k+1), \varphi_i(k)}} = \frac{\mu(\varphi_i(k))}{\mu(\varphi_i(0))}$$

We now assume that C_G is finite, as well as the number of infinite branches ($I = \llbracket 1, \dots, N \rrbracket$ with $N \in \mathbb{N}$), and non empty (the case where C_G is empty corresponds to $G = \mathbb{Z}$, which has been treated in the previous section). With those assumptions, we set:

$$\bar{G} = G \cup \bigcup_{i \in I} \{\Delta_i\}$$

and we define a distance \bar{d} on \bar{G} as follows:

- If $p \in Q_i$, $q \in \bar{Q}_i$, $i \in I$ with $\varphi_i^{-1}(p) \leq \varphi_i^{-1}(q)$:

$$\bar{d}(q, p) = \bar{d}(p, q) = \sum_{n=\varphi_i^{-1}(p)+1}^{\varphi_i^{-1}(q)} \frac{1}{n^2}$$

- If $p, q \in C_G$:

$$\bar{d}(p, q) = \inf\{\text{card}(\gamma), \gamma \text{ path between } p \text{ and } q\} - 1$$

that is, the length of the shortest path (in particular, if p and q are adjacent, then $\bar{d}(p, q) = 1$).

- If $i \in I$ and (p_i, q_i) is the unique adjacent pair in $Q_i \times C_G$, we set:

$$\bar{d}(q, p) = \bar{d}(p, q) := 1 + \bar{d}(p, p_i) + \bar{d}(q_i, q), \quad \forall p \in \bar{Q}_i, \quad \forall q \in C_G$$

and:

$$\bar{d}(q, p) = \bar{d}(p, q) := \bar{d}(p, q_i) + \bar{d}(q_i, q), \quad \forall p \in \bar{Q}_i \quad \forall q \in \bar{Q}_j \quad (i \neq j)$$

Remark 4.2. If $p_1, p_2, p_3 \in Q_i$, $i \in I$, are three vertices of the same infinite branch, then $\bar{d}(p_1, \Delta_i) < \bar{d}(p_2, \Delta_i) < \bar{d}(p_3, \Delta_i)$, if and only if the unique path between p_1 and p_3 contains p_2 .

One easily checks that \bar{G} endowed with this metric is compact: the induced topology on each branch is that of the Alexandrov compactification.

Let G^* be the set of all non empty connected compact subsets of \bar{G} , endowed with the Hausdorff distance:

$$d^*(Q, Q') := \max\left\{\sup_{p \in Q} \inf_{p' \in Q'} \bar{d}(p, p'), \sup_{p' \in Q'} \inf_{p \in Q} \bar{d}(p, p')\right\}$$

It is a closed subspace of the space of compact subsets of \bar{G} (endowed with the Hausdorff distance). We recall that the latter is compact, thus so is G^* .

Let us describe briefly those two metric spaces.
Denoting by $B_d(x, \varepsilon)$ (respectively $\bar{B}_d(x, \varepsilon)$) the open (respectively closed) ball with center x and radius ε for a distance d , we have:

$$\begin{aligned} B_{\bar{d}}(p, 1/2) &= \bar{B}_{\bar{d}}(p, 1/2) = \{p\}, \quad \forall p \in C_G \\ B_{\bar{d}}(\varphi_i(p), \frac{1}{(p+1)^2}) &= \bar{B}_{\bar{d}}(\varphi_i(p), \frac{1}{2(p+1)^2}) = \{\varphi_i(p)\}, \quad \forall p \in \mathbb{N}, \forall i \in I \\ B_{\bar{d}}(\Delta_i, \varepsilon) &= \{\varphi_i(n) : \sum_{j=n+1}^{+\infty} \frac{1}{j^2} < \varepsilon\} \quad \forall i \in I, \forall \varepsilon < \sum_{j=0}^{+\infty} \frac{1}{j^2} \\ \bar{B}_{\bar{d}}(\Delta_i, \varepsilon) &= \{\varphi_i(n) : \sum_{j=n+1}^{+\infty} \frac{1}{j^2} \leq \varepsilon\} \quad \forall i \in I, \forall \varepsilon < \sum_{j=0}^{+\infty} \frac{1}{j^2} \end{aligned}$$

The sets above thus are the elementary open sets for the topology induced by the distance \bar{d} . In particular, each singleton is closed (being the complementary set of a countable number of open sets), and each singleton of G is also open. In addition, the Δ_i are the only limit points of \bar{G} .

The elements of G^* are compact (hence closed) subsets of \bar{G} , that is, finite unions of singletons of G and of neighborhoods of some Δ_i such as described earlier.

The previous statements are easy to check from definition. The following result is not much harder:

Lemma 4.3. *Let $Q \in G^*$, and $I_0 := \{i \in I : \Delta_i \in Q\}$. Then, for all $\eta > 0$, there exists $\varepsilon > 0$ such that for all $Q' \in B_{\bar{d}^*}(Q, \varepsilon)$:*

$$Q' \setminus \left(\bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta) \right) = Q \setminus \left(\bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta) \right)$$

In other words, a "sufficiently close" neighbor of Q must coincide with Q outside the neighborhoods of the $\Delta_i \in G$. In particular, if $Q \in G^*$ is included in G , then Q is an isolated point (in the sense that $\{Q\}$ is open).

Proof. One only needs to take:

$$\varepsilon < \inf \{ \bar{d}(p, q) : p, q \notin \bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta), p \in Q, q \notin Q \}$$

□

For every $i \in I \cup \{0\}$, we set:

$$G_i^* := \{Q \in G^* : |\{j \in I : \Delta_j \in Q\}| = i\} \quad (42)$$

the set of all non empty connected compact subsets of G which contains exactly i points Δ_j .

We define a generator \mathcal{L}^* on G^* by:

$$\forall Q \neq Q' \in G^*, \quad \mathcal{L}_{Q, Q'}^* := \begin{cases} \left(\frac{\mu(Q')}{\mu(Q)} \right) \sum_{q \in Q} \mathcal{L}_{p, q} & \text{if } Q' = Q \cup \{p\}, p \notin Q \\ \left(\frac{\mu(Q')}{\mu(Q)} \right) \sum_{q \notin Q} \mathcal{L}_{p, q} & \text{if } Q' \in \mathcal{C}(Q \setminus \{p\}), p \in Q \\ 0 & \text{otherwise} \end{cases} \quad (43)$$

(we recall that $\mathcal{C}(Q)$ is the set of all connected components of Q).

We observe that the above coefficients are well defined, since if $Q' \in \mathcal{C}(Q \setminus \{p\}) \cap \mathcal{C}(Q \setminus \{q\})$ with $p, q \in Q$, then necessarily $p, q \notin Q'$ and $Q' \cup \{p\}, Q' \cup \{q\}$ are connected, so $p = q$.

The dynamic of a minimal process X^* with generator \mathcal{L}^* should be understood this way: at each jumping time, we add to X^* an adjacent point, or we remove a point that is adjacent to

its complementary set. When X^* disconnects, we choose one of its connected components with a probability proportional to its weight relative to μ .

From this, we deduce that the number of possible transitions is bounded: given $Q \in G^*$, the number of $Q \cup \{p\}$ with $p \in V(Q) \setminus Q$ is bounded by $N + |C_G|$ and the number of $Q' \in \mathcal{C}(Q \setminus \{p\})$, with $p \in V(Q^c) \cap Q$ is bounded (roughly) by $|C_G|^2$.

Lemma 4.4. \mathcal{L}^* is a stratified generator on G^* , and (G_0^*, \dots, G_N^*) is a stratification of \mathcal{L}^* .

Proof. The G_i^* are disjoint and the coefficients of \mathcal{L}^* prevent any transition from some set G_i^* toward some G_j^* , if $j > i$. The first two conditions of Definition 2.4 are then fulfilled.

Let X^* be a trajectory of a minimal process with generator \mathcal{L}^* . We call T_n , $n \in \mathbb{N}$, its jumping times, $(Y_n)_{n \in \mathbb{N}}$ its jump chain and $T = \lim T_n$ its explosion time, that we assume to be finite. Thus Y passes through infinitely many states, and visits each of them finitely many times. Since G^* is compact, it follows that Y has at least one limit point, which is a limit point of G^* . If it is unique, then Y converges to this point. Let us assume by contradiction that there are two distinct limit points, that we call E_1 and E_2 , respective limits of two subsequences $(Y_{\phi_1(n)})_n$ and $(Y_{\phi_2(n)})_n$ of $(Y_n)_n$. The idea of what follows is to show that we cannot go infinitely many times from one subsequence to the other one.

Since E_1 and E_2 are two connected compact subsets of \bar{G} and $E_1 \neq E_2$, there exists $p \in G$ such that $p \in E_1$ and $p \notin E_2$ (or the other way round).

From Lemma 4.3, there exist $n_1, n_2 \in \mathbb{N}$ such that $p \in Y_{\phi_1(n)}$, $\forall n > n_1$ and $p \notin Y_{\phi_2(n)}$, $\forall n > n_2$. After the time $\max\{\phi_1(n_1), \phi_2(n_2)\}$ the chain Y jumps infinitely many times from a state Q containing an adjacent point of p toward the state $Q \cup \{p\}$. This is impossible almost surely, the rate of such a transition being upper bounded by:

$$\left(1 + \sup_{q \in V(p)} \frac{\mu(p)}{\mu(q)} \sum_{q' \in V(p)} \mathcal{L}_{p,q'}\right) < \infty$$

Thus, the sequence $(Y_n)_n$ has an almost surely unique limit $E \in G_k^*$, $k \in I \cup \{0\}$. Let $I_0 := \{i \in I : \Delta_i \in E\}$. From Lemma 4.3, for all $\eta > 0$, there exists n_1 such that for every $n > n_1$:

$$Y_n \setminus \left(\bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta) \right) = E \setminus \left(\bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta) \right)$$

In particular, if $i \notin I_0$, then $\Delta_i \notin Y_n$. In addition, there exists $n > n_1$ such that $Y_n \neq E$, and since Y_n is connected, there exists $i \in I_0$ such that $\Delta_i \notin Y_n$. There are therefore less Δ_i in Y_n than in E , that is, $Y_n \in G_j^*$ with $j < k$.

Now, the number of transitions of the minimal process from some state in G_i^* toward some other one in G_j^* , $i \neq j$, is finite a.s. (since they are only possible if $j < i$) and thus there exists a last one before the explosion. Let T_{n_2} be the time of this last transition, and G_j^* , $j < k$, the arrival set $(Y_{n_2} \in G_j^*)$, then, for all $t \in [T_{n_2}, T[$, $X_t^* \in G_j^*$. So the third condition is also satisfied. \square

We define the kernel $\Lambda(\cdot, \cdot)$ by:

$$\forall Q \in G^*, p \in G, \quad \Lambda(Q, p) = \frac{\mu(p)\delta_p(Q)}{\mu(Q)}$$

Λ is continuous in the first variable, and:

Lemma 4.5. \mathcal{L} and \mathcal{L}^* satisfy $\mathcal{L}^* \Lambda = \Lambda \mathcal{L}$.

Proof. Let $Q \in G^*$ and $p \in G$. If $p \notin Q$, then:

$$\begin{aligned}
\sum_{Q' \in G^*} \mathcal{L}_{Q,Q'}^* \Lambda(Q', p) &= \sum_{q \in G} \mathcal{L}_{Q, Q \cup \{q\}}^* \frac{\mu(p) \delta_p(Q \cup q)}{\mu(Q \cup \{q\})} \\
&= \mathcal{L}_{Q, Q \cup \{p\}}^* \frac{\mu(p)}{\mu(Q \cup \{p\})} \\
&= \frac{\mu(p)}{\mu(Q)} \sum_{q \in Q} \mathcal{L}_{p,q} \\
&= \sum_{q \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q,p} \\
&= \sum_{q \in Q} \Lambda(Q, q) \mathcal{L}_{q,p}
\end{aligned}$$

On the other hand, if $p \in Q$, then:

$$\begin{aligned}
\sum_{Q' \in G^*} \mathcal{L}_{Q,Q'}^* \Lambda(Q', p) &= \sum_{q \notin Q} \mathcal{L}_{Q, Q \cup \{q\}}^* \frac{\mu(p)}{\mu(Q \cup \{q\})} + \mathcal{L}_{Q,Q}^* \Lambda(Q, p) \\
&\quad + \sum_{q \neq p \in Q} \sum_{Q' \in \mathcal{C}(Q \setminus \{q\})} \mathcal{L}_{Q,Q'}^* \frac{\mu(p)}{\mu(Q \setminus \{q\})} \\
&= \sum_{q \notin Q} \sum_{q' \in Q} \frac{\mu(p)}{\mu(Q)} \mathcal{L}_{q,q'} + \sum_{q \neq p \in Q} \sum_{q' \notin Q} \frac{\mu(p)}{\mu(Q)} \mathcal{L}_{q,q'} \\
&\quad + \frac{\mu(p)}{\mu(Q)} \left(\sum_{q \notin Q} \sum_{q' \notin Q} \mathcal{L}_{q,q'} + \sum_{q \in Q} \sum_{q' \in Q} \mathcal{L}_{q,q'} \right) \\
&= \frac{\mu(p)}{\mu(Q)} \left(\sum_{q \notin Q, q' \in G} \mathcal{L}_{q,q'} + \sum_{q \neq p \in Q, q' \in G} \mathcal{L}_{q,q'} + \sum_{q' \in Q} \mathcal{L}_{p,q'} \right) \\
&= \sum_{q \in Q} \frac{\mu(p)}{\mu(Q)} \mathcal{L}_{p,q} \\
&= \sum_{q \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q,p}
\end{aligned}$$

We then have proved that $(\mathcal{L}^* \Lambda)_{Q,p} = (\Lambda \mathcal{L})_{Q,p}$. □

Thanks to the previous two lemmas and Theorem 2.10 we deduce that whatever the initial distribution μ_0 of X , there exists a stratified jump process X^* with generator \mathcal{L}^* , which is Λ -linked to X (taking for an initial distribution of X^* , for example, $\mu_0^*(A) = \sum_{x \in S} \mu_0(x) \mathbf{1}_{A=\{x\}}$ for all $A \in \mathcal{S}^*$). In that case, by definition of Λ , X^* is also a strong stationary dual of X .

One will have noticed that a process X^* with generator \mathcal{L}^* has $N+1$ absorbing states: each of the singletons $\{\Delta_i\}$ and \bar{G} . However, by intertwining X and X^* with Λ , we get:

$$\begin{aligned}
0 &= \mathbb{P}(X_t = \Delta_i \mid X_0^* \neq \{\Delta_i\}) \\
&\geq \mathbb{P}(X_t = \Delta_i \mid X_t^* = \{\Delta_i\}, X_0^* \neq \{\Delta_i\}) \mathbb{P}(X_t^* = \{\Delta_i\} \mid X_0^* \neq \{\Delta_i\}) \\
&= \mathbb{P}(X_t = \Delta_i \mid X_t^* = \{\Delta_i\}) \mathbb{P}(X_t^* = \{\Delta_i\} \mid X_0^* \neq \{\Delta_i\}) \\
&= \mathbb{P}(X_t^* = \{\Delta_i\} \mid X_0^* \neq \{\Delta_i\})
\end{aligned} \tag{44}$$

since X is nonexplosive. We conclude that the only absorbing state that may be reached in finite time is \bar{G} .

It remains to study the time to absorption of X^* . More precisely, we will study the behavior of the bounds of X^* on each infinite branch, by comparing them with birth-death processes on \mathbb{N} .

For every $i \in I$, and every $p \in \mathbb{N}$, we set:

$$G_p^i := \bar{G} \setminus \varphi_i(\llbracket p+1, \infty \rrbracket) = \bigcup_{j \neq i} \bar{Q}_j \cup C_G \cup \varphi_i(\llbracket 0, p \rrbracket) \in G^* \quad (45)$$

(that is, the point of G^* containing all \bar{G} except the points of \bar{Q}_i beyond $\varphi_i(p)$), and:

$$\begin{cases} \mathcal{L}_{p,p+1}^i := \left(\frac{\mu(G_{p+1}^i)}{\mu(G_p^i)} \right) \mathcal{L}_{\varphi_i(p+1), \varphi_i(p)} = \mathcal{L}_{G_p^i, G_{p+1}^i}^* \\ \mathcal{L}_{p+1,p}^i := \left(\frac{\mu(G_p^i)}{\mu(G_{p+1}^i)} \right) \mathcal{L}_{\varphi_i(p+1), \varphi_i(p+2)} = \mathcal{L}_{G_{p+1}^i, G_p^i}^* \\ \mathcal{L}_{p,p}^i = -\mathcal{L}_{p,p+1}^i - \mathcal{L}_{p,p-1}^i \end{cases}$$

which define generators $(\mathcal{L}_{p,q}^i)_{p,q \in \mathbb{N}}$ of birth-death processes on \mathbb{N} .

Lemma 4.6. *There exists $\lambda > 0$ such that, for every $i \in I$ and every probability measure μ_0^* on G^* , there exists a family (X^*, Y, Z^i) such that:*

- The couple (X^*, Z^i) is Markovian, and its marginals X^* and Z^i are Markov processes, with respective generators \mathcal{L}^* and \mathcal{L}^i .
- Y is an exponential random variable with parameter λ , independent from Z^i .
- X^* has initial distribution μ_0^* .
- For every $t > s > 0$:

$$\mathbb{P}(X_t^* \cap \varphi_i(\llbracket Z_t^i, \infty \rrbracket) \neq \emptyset \mid t < Y, X_0^* \not\subset G_0^i) = 1$$

$$\mathbb{P}(X_s^* \cap C_G \subset X_t^* \mid t < Y) = 1$$

Proof. We are first going to distinguish different kinds of transition for the process X^* , when it is at a state $Q \in G^*$ such that $Q \cap Q_i \neq \emptyset$. The first kind is that which changes X^* only by a point of Q_i . It concerns the transitions from Q to $Q \cup \{\varphi_i(p)\}$, or to $Q \setminus \{\varphi_i(p-1)\}$, where $p = \sup\{q \in \mathbb{N} : \varphi_i(q) \in Q\} + 1$. These transitions "make us one step closer to or further from" a state containing the point Δ_i . We will thus denote $A_i^+(Q) = Q \cup \{\varphi_i(p)\}$, $A_i^-(Q) = Q \setminus \{\varphi_i(p-1)\}$ and $A_i(Q) = \{A_i^+(Q), A_i^-(Q)\}$.

The second kind of transitions may remove several points at once. To do so, there must exist $q \in Q \in G^*$ such that $Q \setminus \{q\}$ is disconnected and $q \in V(Q^c)$. This is only possible if $q \in C_G$. Thus, we naturally set:

$$B(Q) := \{Q' \in \mathcal{C}(Q \setminus \{q\}) : q \in Q \cap C_G\}$$

The third kind gathers all other transitions, those which do not interfere in our problematic:

$$C_i(Q) = \{Q' \in G^* : \mathcal{L}_{Q,Q'} > 0, Q' \notin A_i(Q) \cup B(Q)\}$$

Those transitions remove a point of Q_j , $j \neq i$ or add a point different from $\varphi_i(p)$.

We will talk about a A_i^+ (resp. A_i^- , resp. B , resp. C) transition for a transition from Q to some $Q' = A_i^+(Q)$ (resp $Q' = A_i^-$, resp. $Q' \in B(Q)$, resp. $Q' \in C_i(Q)$).

We observe that the set $\{\mathcal{L}_{Q,Q'}^* : Q \in G^*, Q' \in B(Q)\}$ is bounded. Indeed, if $Q \in G^*$ is such that $B(Q)$ is non empty, then it contains a point of C_G , so:

$$\mu(Q) \geq \inf\{\mu(p) : p \in C_G\} =: a_1$$

and if $p_0 \in Q \cap C_G$, then:

$$\sum_{q \notin Q} \mathcal{L}_{q,p_0} \leq \sup_{p \in C_G} \sum_{q \in V(C_G)} \mathcal{L}_{q,p} =: a_2$$

which gives us for all $Q' \in \mathcal{C}(Q \setminus \{p_0\})$:

$$\begin{aligned}\mathcal{L}_{Q,Q'}^* &= \frac{\mu(Q')}{\mu(Q)} \sum_{q \notin Q} \mathcal{L}_{q,p_0} \\ &\leq \frac{1}{a_1} a_2\end{aligned}$$

The set $\{|B(Q)| : Q \in G^*\}$ is also bounded by a constant a_3 (roughly, $N + |C_G|$), so we deduce:

$$\sum_{Q' \in B(Q)} \mathcal{L}_{Q,Q'}^* \leq \frac{a_2 a_3}{a_1} =: \lambda$$

and λ does not depend on Q .

We finally observe that for every $p \in \mathbb{N}$, if $\varphi_i(p) \in Q$ and $\varphi_i(p+1) \notin Q$, $Q \in G^*$, then $\mathcal{L}_{Q,A_i^+(Q)}^* \geq \mathcal{L}_{p,p+1}^i$ and $\mathcal{L}_{Q,A_i^-(Q)}^* \leq \mathcal{L}_{p,p-1}^i$.

We are now going, thanks to the above observations, to construct the family (X^*, Y, Z^i) by preventing any B transition before time Y , and by forcing Z^i to go down in case of a A_i^- transition or by setting a A_i^+ transition if Z^i goes up, when necessary. Of course, this construction is only made when X_0^* contains at least one point of Q_i .

Construction of the family (X^*, Y, Z^i) : In order to be reduced to a Markov process, we will construct instead (X^*, N, Z^i) , where N is a Poisson process with intensity λ . Y is then the first jumping time of N . Firstly, we draw X_0^* with distribution μ_0^* . If $X_0^* \cap Q_i = \emptyset$, we can stop the construction. Otherwise, we set $Z_0^i = \sup\{p \in \mathbb{N} : \varphi_i(p) \in X_0^*\}$ and $T_0 = 0$.

We assume (X^*, N, Z^i) to be constructed up to the time T_n , $n \geq 0$, $Z_{T_n}^i = p$ and $X_{T_n}^* = Q$. If $N_t \geq 1$, the construction is over and we let X^* and Z evolve independently. Otherwise, we separate two cases:

- *First case:* $\varphi_i(Z_{T_n}^i) \notin A_i^-(X_{T_n}^*)$

This is the case where $Z_{T_n}^i$ is "at the far end" of $X_{T_n}^*$, and at this point we must be careful to couple Z^i and X^* so that Z^i stays "inside" at time T_{n+1} . We thus draw an exponential variable ϵ_n^1 with parameter λ and another one, ϵ_n^2 , independent from ϵ_n^1 and with parameter:

$$\Lambda_n^2 := \sum_{Q' \in C_i(Q)} \mathcal{L}_{Q,Q'}^* + \mathcal{L}_{Q,A_i^+(Q)}^* + \mathcal{L}_{p,p-1}^i$$

We set $T_{n+1} = T_n + \epsilon_n^1 \wedge \epsilon_n^2$. If $\epsilon_n^1 < \epsilon_n^2$, we enforce $N_{T_{n+1}} = 1$ and choose $X_{T_{n+1}}^* = Q'$ with probability $\frac{\mathcal{L}_{Q,Q'}^*}{\Lambda_n^2}$, for all $Q' \in B(Q)$. The sum of these probabilities may be less than 1, so if no state is chosen we stay in Q .

If $\epsilon_n^1 > \epsilon_n^2$, we leave $N_{T_{n+1}}$ at 0, and we set:

$$(X_{T_{n+1}}^*, Z_{T_{n+1}}^i) = \begin{cases} (Q', p) & \text{with probability } \frac{\mathcal{L}_{Q,Q'}^*}{\Lambda_n^2}, Q' \in C_i(Q) \\ (Q', p+1) & \text{with probability } \frac{\mathcal{L}_{p,p+1}^i}{\Lambda_n^2}, Q' = A_i^+(Q) \\ (Q', p) & \text{with probability } \frac{\mathcal{L}_{Q,Q'}^* - \mathcal{L}_{p,p+1}^i}{\Lambda_n^2}, Q' = A_i^+(Q) \\ (Q, p-1) & \text{with probability } \frac{\mathcal{L}_{p,p-1}^i - \mathcal{L}_{Q,A_i^-(Q)}^*}{\Lambda_n^2} \\ (Q', p-1) & \text{with probability } \frac{\mathcal{L}_{Q,A_i^-(Q)}^*}{\Lambda_n^2}, Q' = A_i^-(Q) \end{cases}$$

- *Second case:* $\varphi_i(Z_{T_n}^i) \in A_i^-(X_{T_n}^*)$

In that case, we do not need to ensure that "the far end" of X^* stays above Z^i , but only to check if a B transition occurs. Consequently, we draw an exponential variable ϵ_n^1 with parameter λ and another one, ϵ_n^2 , independent from ϵ_n^1 and with parameter:

$$\Lambda_n^2 := \sum_{Q' \in C_i(Q) \cup A_i(Q)} \mathcal{L}_{Q,Q'}^* + \mathcal{L}_{p,p+1}^i + \mathcal{L}_{p,p-1}^i$$

We set $T_{n+1} = T_n + \epsilon_n^1 \wedge \epsilon_n^2$. If $\epsilon_n^1 < \epsilon_n^2$, we proceed exactly as in the first case. Otherwise, we chose:

$$(X_{T_{n+1}}^*, Z_{T_{n+1}}^i) = \begin{cases} (Q', p) & \text{with probability } \frac{\mathcal{L}_{Q,Q'}^*}{\Lambda_n^2}, \quad Q' \in C_i(Q) \cup A_i(Q) \\ (Q, p+1) & \text{with probability } \frac{\mathcal{L}_{p,p+1}^i}{\Lambda_n^2} \\ (Q, p-1) & \text{with probability } \frac{\mathcal{L}_{p,p-1}^i}{\Lambda_n^2} \end{cases}$$

By construction, we do have $X_t^* \cap \varphi_i(\llbracket Z_t^i, \infty \rrbracket) \neq \emptyset$ if $t < Y$, and Y and Z^i independent, until the first explosion time $\tau_1 := \lim T_n$. One checks that the marginal processes have the desired generators. Starting from the explosion time, we carry on the construction the same way.

□

Lemma 4.7. *For all $i \in I$ if Z^i is a process with generator \mathcal{L}^i , then the explosion time of Z_t^i is finite almost surely if and only if the condition (11):*

$$\sum_{j=1}^{\infty} \mu^i(j+1) \sum_{k=1}^j \frac{1}{\mu^i(k) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} < \infty$$

holds. In that case, Z^i goes to $+\infty$ in finite time, otherwise it is nonexplosive.

Proof. We recall that the explosion criterion for Z^i is:

$$\sum_{i=1}^{\infty} \prod_{j=1}^i \frac{\mathcal{L}_{j,j-1}^i}{\mathcal{L}_{j,j+1}^i} \sum_{k=1}^i \prod_{l=1}^k \frac{\mathcal{L}_{j-1,j}^i}{\mathcal{L}_{j,j-1}^i} < \infty$$

(cf [2], Section 8.1). That is:

$$\begin{aligned} & \infty > \sum_{n=1}^{\infty} \prod_{j=1}^n \frac{\left(\frac{\mu(G_{j-1}^i)}{\mu(G_j^i)} \right) \mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}}{\left(\frac{\mu(G_{j+1}^i)}{\mu(G_j^i)} \right) \mathcal{L}_{\varphi_i(j+1), \varphi_i(j)}} \sum_{k=1}^n \prod_{l=1}^k \frac{\left(\frac{\mu(G_j^i)}{\mu(G_{j-1}^i)} \right) \mathcal{L}_{\varphi_i(j), \varphi_i(j-1)}}{\left(\frac{\mu(G_{j-1}^i)}{\mu(G_j^i)} \right) \mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}} \\ & \geq \sum_{n=1}^{\infty} \prod_{j=1}^n \frac{\mu(G_{j-1}^i) \mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}}{\mu(G_{j+1}^i) \mathcal{L}_{\varphi_i(j+1), \varphi_i(j)}} \sum_{k=1}^n \prod_{l=1}^k \frac{\mathcal{L}_{\varphi_i(j), \varphi_i(j-1)}}{\mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}} \\ & \geq \sum_{n=1}^{\infty} \prod_{j=1}^n \frac{\mu(G_0^i) \mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}}{\mathcal{L}_{\varphi_i(j+1), \varphi_i(j)}} \sum_{k=1}^n \prod_{l=1}^k \frac{\mathcal{L}_{\varphi_i(j), \varphi_i(j-1)}}{\mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}} \\ & \geq \frac{\mu(G_0^i)}{\mu(\varphi_i(0))} \sum_{n=1}^{\infty} \mu(\varphi_i(n+1)) \sum_{k=1}^n \frac{1}{\mu(\varphi_i(k)) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} \\ & = \frac{\mu(G_0^i)}{\mu(\varphi_i(0))} \sum_{n=1}^{\infty} \mu^i(n+1) \sum_{k=1}^n \frac{1}{\mu^i(k) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} \end{aligned}$$

where the G_j^i are the subsets of \bar{G} defined in (45).

□

The purpose of these generators \mathcal{L}^i is then to provide a sufficient condition for the explosion of a process X^* with generator \mathcal{L}^* along each branch Q_i . We get a necessary condition similarly, defining, for every $i \in I$ and every $n \in \mathbb{N}$ and $p \geq n$:

$$\begin{cases} \mathcal{L}_{p,p+1}^{i,n} := \left(1 + \frac{\mu(\varphi_i(p+1))}{\mu(\varphi_i([n,p]))}\right) \mathcal{L}_{\varphi_i(p+1), \varphi_i(p)} = \mathcal{L}_{\varphi_i([n,p]), \varphi_i([n,p+1])}^* \\ \mathcal{L}_{p,p-1}^{i,n} := \left(1 - \frac{\mu(\varphi_i(p))}{\mu(\varphi_i([n,p]))}\right) \mathcal{L}_{\varphi_i(p-1), \varphi_i(p)} = \mathcal{L}_{\varphi_i([n,p]), \varphi_i([n,p-1])}^* \\ \mathcal{L}_{p,p}^{i,n} = -\mathcal{L}_{p,p+1}^{i,n} - \mathcal{L}_{p,p-1}^{i,n} \end{cases}$$

The generators $(\mathcal{L}_{p,q}^{i,n})_{p,q \geq n}$ define birth-death processes on $[n, +\infty]$. Exactly as above, we get the following lemmas:

Lemma 4.8. *For every $i \in I$, and every probability measure μ_0^* on G^* , there exists a process X^* with generator \mathcal{L}^* and initial distribution μ_0^* and a sequence $(Z^{i,n})_{n \in \mathbb{N}}$ of birth death processes such that, for every $n \in \mathbb{N}$, $Z^{i,n}$ has generator $\mathcal{L}^{i,n}$ and:*

$$\mathbb{P}(X_t^* \cap \varphi_i(\llbracket Z_t^{i,n}, \infty \rrbracket) \subset \{\varphi_i(Z_t^{i,n})\} \mid \Delta_i \notin X_0^*, t < T_n) = 1$$

for all $t > 0$, with $T_n := \inf\{s > 0 : X_s^* \cap G_n^i = \emptyset\}$.

Proof. We only need to adapt slightly the proof of Lemma 3.3. This time, we let X^* and each of the $Z^{i,n}$ evolve independently up to the time $\tau := \inf\{t > 0 : \exists n \in \mathbb{N}, Z^{i,n} \in X_t^*\}$, then we apply the same method to force X^* to stay in $G_{Z^{i,n}}^i$. \square

Lemma 4.9. *If $Z^{i,n}$ is a process with generator $\mathcal{L}^{i,n}$, with $i \in I$, $n \in \mathbb{N}$ then the explosion time of $Z^{i,n}$ is finite almost surely if and only if:*

$$\sum_{j=n+1}^{\infty} \mu^i(j+1) \sum_{k=n+1}^j \frac{1}{\mu^i(k) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} < \infty \quad (46)$$

Theorem 4.10. *Let X^* be a process with generator \mathcal{L}^* . If there exists $i \in I$ such that the condition (11) is not fulfilled, then for all $t > 0$:*

$$\mathbb{P}(\Delta_i \in X_t^* \mid \Delta_i \notin X_0^*) = 0$$

Conversely, if for every $i \in I$ the condition (11) is fulfilled then, almost surely, there exists $T < \infty$ such that $X_t^ = \bar{G}$, $\forall t > T$.*

Proof. We prove the first part with Lemmas 4.8 and 4.9. Let $i \in I$ such that the condition (11) is not fulfilled. We see immediately that for every $n \in \mathbb{N}$, the condition (46) is not either. Let then $(\tilde{X}^*, (Z^{i,n})_{n \in \mathbb{N}})$ be a family of processes as in Lemma 4.8, \tilde{X}^* having same law as X^* . We then get, for all $t > 0$:

$$\begin{aligned} \mathbb{P}(\Delta_i \in X_t^* \mid \Delta_i \notin X_0^*) &= \mathbb{P}(\Delta_i \in \tilde{X}_t^* \mid \Delta_i \notin \tilde{X}_0^*) \\ &\leq \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} \{\tilde{X}_t^* \cap \varphi_i(\llbracket Z_t^{i,n} + 1, \infty \rrbracket) \neq \emptyset\} \mid \Delta_i \notin \tilde{X}_0^*\right) \\ &\leq \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} \{t \geq T_n\} \mid \Delta_i \notin \tilde{X}_0^*\right) \\ &= \mathbb{P}(\exists s \leq t : \tilde{X}_s^* = \{\Delta_i\} \mid \Delta_i \notin \tilde{X}_0^*) \\ &= \mathbb{P}(\tilde{X}_t^* = \{\Delta_i\} \mid \Delta_i \notin \tilde{X}_0^*) \end{aligned}$$

and the latter is equal to 0 from (44).

We show the second part with Lemmas 4.6 and 4.7. Observe that since X^* has the law of a process Λ -linked to X , which is positive recurrent, X^* must reach C_G in finite time almost surely. Thus we may assume without loss of generality that the process starts from this moment, so that X_0^* contains at least one point of C_G .

When the process X^* reaches a state containing some Δ_i (necessarily after an explosion), the latter can exit X^* only in the same time as all the rest of the branch Q_i and the point p_i of C_G adjacent to Q_i . This event corresponds to a B transition. The idea is therefore to show that the process can reach successively each of the Δ_i without losing any point of C_G (so without losing the Δ_i already reached). To do so, we find a lower bound, uniform on the initial condition, of the probabilities to reach each Δ_i before a B transition. We do this in two steps: first, we find a lower bound of the probability to reach $\varphi_i(0)$ by following a given path, then we couple the process with Y and Z^i as in Lemma 4.6.

For every $i \in I$, we call $f_i : G^* \rightarrow G^*$ the function defined by:

$$f_i(Q) = Q \cap (C_G \cup \{\varphi_i(0)\})$$

A way to ensure the absence of B transition on a time interval $[0, t]$ is to ask for $f_i(X^*)$ to be increasing (for set inclusion) on this interval.

Let $t > 0$, $i \in I$, $C \subset C_G \cup \{\varphi_i(0)\}$ such that $C \not\subset \{\varphi_i(0)\}$ (so C contains at least one point of C_G), $Q \in G^*$ such that $f_i(Q) = C$ and (q_0, \dots, q_{n_0}) a path with $q_0 \in C$, $q_1, \dots, q_{n_0-1} \in C_G \setminus C$ and $q_{n_0} = \varphi_i(0)$. If $n_0 = 0$, that means that $\varphi_i(0) \in Q$. We write $C^n = C \cup \{q_1\} \cup \dots \cup \{q_n\}$, $0 \leq n \leq n_0$. We also write T_k for the k -th jumping time of X^* , $k \in \mathbb{N}$, with $T_0 = 0$, $(Y_k)_{k \in \mathbb{N}}$ the underlying jump chain: $X_{T_k}^* = Y_k$ for every $k \in \mathbb{N}$ and $T_{\phi(k)}$ the subsequence of jumping times of $f_i(X^*)$. The processes $(T_{\phi(k)})$ and $(f_i(Y_{\phi(k)}))$, expressing the evolution of X^* on $C_G \cup \{\varphi_i(0)\}$, are not Markovian, but we can bound their transition probabilities using the process X^* . The first aim is to find a lower bound of $\mathbb{P}(f_i(X_t^*) = C^{n_0}, f_i(X^*) \text{ increasing on } [0, t] \mid X_0^* = Q)$ that does not depend on Q . We start roughly:

$$\begin{aligned} & \mathbb{P}(f_i(X_t^*) = C^{n_0}, f_i(X^*) \text{ increasing on } [0, t] \mid X_0^* = Q) \\ & \geq \mathbb{P}\left(\bigcap_{k=1}^{n_0} \{T_{\phi(k)} - T_{\phi(k-1)} \leq t/n_0, f_i(Y_{\phi(k)}) = C^k, T_{\phi(n_0+1)} - T_{\phi(n_0)} > t\} \mid X_0^* = Q\right) \\ & = \prod_{k=1}^{n_0} \mathbb{P}(\epsilon_k \in [0, t/n_0]) p_k \times \mathbb{P}(\epsilon_{n_0+1} > t/n_0) \end{aligned} \quad (47)$$

where for every $k \leq n_0 + 1$:

$$\begin{aligned} \epsilon_k & \sim \mathcal{L}(T_{\phi(k)} - T_{\phi(k-1)} \mid \bigcap_{n=1}^{k-1} \{f_i(Y_{\phi(n)}) = C^n, T_{\phi(n)} - T_{\phi(n-1)} \in [0, t/n_0], X_0^* = Q\}) \\ p_k & = \mathbb{P}(f_i(Y_{\phi(k)}) = C^k \mid \bigcap_{n=0}^{k-1} \{f_i(Y_{\phi(n)}) = C^n, T_{\phi(n+1)} - T_{\phi(n)} \in [0, t/n_0], X_0^* = Q\}) \end{aligned}$$

(observe that in the above conditional probability, $f_i(Y_0) = C^0$ is not necessary since it is implied by $X_0^* = Q$, and was left only by simplicity of notation).

If at time $T_{\phi(k)}$ the process X^* goes from some state Q' to Q'' then we have $Q'' = Q' \cup \{p\}$, $p \in f_i(V(Q'))$ or $Q'' \in \mathcal{C}(Q' \setminus \{p\})$, $p \in f_i(Q')$, and one easily checks from the coefficients of \mathcal{L}^* that:

$$\left(\min_{q \in V(f_i(G))} \mu(q)\right) \left(\min_{\substack{q \in V(C_G) \\ q' \in V(q)}} \mathcal{L}_{q,q'}\right) \leq \mathcal{L}_{Q',Q''}^* \leq \frac{\max_{q \in V(C_G)} |\mathcal{L}_{q,q}|}{\min_{q' \in C_G} \mu(q')}$$

We call $\alpha_1 < \infty$ the r.h.s. and $\alpha_2 > 0$ the l.h.s. in the above inequality. One observes that for a given Q' , the number of possible Q'' of the form $Q' \cup \{p\}$ is upper bounded by $|C_G|$, and $|\bigcup_{p \in f_i(Q')} \mathcal{C}(Q' \setminus \{p\})|$ is upper bounded by $(|C_G| + 1)^2$. Conditioning by $X_{T_{\phi(k)}-1}^*$, we get:

$$\begin{aligned}
p_k &\geq \inf \{ \mathbb{P}(X_{T_{\phi(k)}}^* = Q' \cup \{p\} \mid X_{T_{\phi(k)-1}}^* = Q'), f_i(Q') \neq \emptyset, p \in V(Q') \cap V(C_G) \} \\
&= \inf \left\{ \frac{\mathcal{L}_{Q', Q' \cup \{p\}}^*}{\sum_{\substack{Q'' \in G^* \\ f_i(Q'') \neq f_i(Q')}} \mathcal{L}_{Q', Q''}^*}, f_i(Q') \neq \emptyset, p \in V(Q') \cap V(C_G) \right\} \\
&\geq \frac{\alpha_2}{(|C_G| + (|C_G| + 1)^2)\alpha_1} =: \alpha_0 > 0
\end{aligned}$$

and, given $X_{T_{\phi(k-1)}}^*$ we can bound $T_{\phi(k)} - T_{\phi(k-1)}$ by two exponential variables E_1 (lower bound) and E_2 (upper bound), with respective parameters:

$$\begin{aligned}
\sup \left\{ \sum_{\substack{Q'' \in G^* \\ f_i(Q'') \neq f_i(Q')}} \mathcal{L}_{Q', Q''}^* : Q' \in G^*, f_i(Q') \neq \emptyset \right\} &\leq (|C_G| + (|C_G| + 1)^2)\alpha_1 \\
\inf \left\{ \sum_{\substack{Q'' \in G^* \\ f_i(Q'') \neq f_i(Q')}} \mathcal{L}_{Q', Q''}^* : Q' \in G^*, f_i(Q') \neq \emptyset \right\} &\geq \alpha_2
\end{aligned}$$

Put in other words:

$$\mathbb{P}(\epsilon_k \in [0, t/n_0]) \geq \mathbb{P}(E_2 \in [0, t/n_0]) \geq 1 - \exp(-\alpha_2 t/n_0)$$

and:

$$\mathbb{P}(\epsilon_{n_0+1} > t) \geq \mathbb{P}(E_1 > t) \geq \exp(-(|C_G| + (|C_G| + 1)^2)\alpha_1 t)$$

Going back to (47) and using the fact that $n_0 \leq |C_G|$, we see that we have found a lower bound ε_t on the probability to reach $\varphi_i(0)$ in time t before any B transition, and that ε_t does not depend on i , nor on the initial configuration C . Hence we have made the first step of the bounding.

Starting from the state reached during this first step, we construct (X^*, Y, Z^i) as in Lemma 4.6. The probability to reach a state containing Δ_i before a B transition, in time t , is lower bounded by $\mathbb{P}(Z_t^i = \infty, Y > t \mid Z_0^i = \varphi_i(0))$ from the law of the couple (X^*, Z^i) , and this probability is positive since Z^i and Y are independent. Doing the two steps consecutively (each one in time $t/2$ for instance), we get:

$$\mathbb{P}(\Delta_i \in X_t^*, f_i(X^*) \text{ increasing on } [0, t] \mid X_0^*, X_0^* \cap C_G \neq \emptyset) \geq \varepsilon_t^i$$

and this constant only depends on i and t . By Markovianity and homogeneity, it follows that:

$$\begin{aligned}
&\mathbb{P}(\Delta_i \in X_{Nt}^*, \forall i \in I \mid X_0^* \cap C_G \neq \emptyset) \\
&\geq \prod_{i=1}^N \inf_{\substack{Q \in G^* \\ Q \cap C_G \neq \emptyset}} \mathbb{P}(\Delta_i \in X_{it}^*, f_i(X^*) \text{ increasing on } [(i-1)t, it] \mid X_{(i-1)t}^* = Q) \\
&\geq \prod_{i=1}^N \inf_{\substack{Q \in G^* \\ Q \cap C_G \neq \emptyset}} \mathbb{P}(\Delta_i \in X_t^*, f_i(X^*) \text{ increasing on } [0, t] \mid X_0^* = Q) \\
&\geq \prod_{i \in I} \varepsilon_t^i > 0
\end{aligned}$$

The set G_N^* of states containing all Δ_i (defined in (42)) is then positive recurrent. It remains to conclude by observing that from some state $Q \in G_N^*$, we can go to \bar{G} in finitely many jumps: if q_1, \dots, q_{n_0} are such that $\{q_i, i \leq n_0\} = C_G \setminus Q$ and $q_i \in V(Q \cup \{q_1\} \cup \dots \cup \{q_{i-1}\})$, then

$$\mathbb{P}(X_{T_i}^* = Q \cup \{q_1\} \cup \dots \cup \{q_i\}, \forall i \leq n_0 \mid X_0^* = Q) > \varepsilon$$

where ε is a positive constant independent from Q and n_0 .

\bar{G} is thus positive recurrent, and the process X^* is absorbed in finite time almost surely. \square

After a last technical lemma, we will finally be able to prove Theorem 1.1:

Lemma 4.11. *If X^* is a stratified jump process with generator \mathcal{L}^* , such that $\mathbb{E}\left(\frac{1}{\mu(X_0^*)}\right) < +\infty$, then $\mathbb{E}\left(\frac{1}{\mu(X_t^*)}\right) < +\infty$ for all $t > 0$.*

Proof. Let $A_n := \bigcup_{i \in I} \varphi_i(\llbracket n, \infty \rrbracket)$ and

$$\tau_n := \inf\{t > 0 : X_t^* \subset A_n\}$$

Observe that the sequence $(A_n)_{n \in \mathbb{N}}$ is decreasing and goes to $A_\infty = \{\Delta_i, i \in I\}$, whereas the sequence $(\tau_n)_{n \in \mathbb{N}}$ is increasing (starting from the first rank n such that $\tau_n \neq 0$) and goes to infinity (since the set A_∞ cannot be reached in finite time). In particular, for all $t > 0$:

$$\mathbb{P}(X_{t \wedge \tau_n}^* \not\subset A_{n+1} \mid X_0^* \not\subset A_{n+1}) = 1$$

In addition, if $X_0^* \not\subset A_{n+1}$, then for every $n \in \mathbb{N}$, we get:

$$\frac{1}{\mu(X_{t \wedge \tau_n}^*)} \leq \sup_{q \in G \setminus A_{n+1}} \frac{1}{\mu(q)} =: K_n$$

We then set, for all $Q \in G^*$ and all $n \in \mathbb{N}$:

$$g_n(Q) = \frac{1}{\mu(Q)} \wedge K_n$$

These functions are continuous and bounded and for every n , g_n coincides with $\frac{1}{\mu}$ on $\{Q \in G^* : Q \not\subset A_{n+1}\}$. Using the fact that

$$Q \not\subset A_n \Rightarrow Q' \not\subset A_{n+1}, \forall Q' \in G^* \text{ such that } \mathcal{L}_{Q,Q'}^* > 0$$

we get, for every $n \in \mathbb{N}$, for all $Q \not\subset A_n$:

$$\begin{aligned} \mathcal{L}^* \frac{1}{\mu}(Q) &= \mathcal{L}^* g_n(Q) \\ &= \sum_{Q' \neq Q} \mathcal{L}_{Q,Q'}^* \left(\frac{1}{\mu(Q')} - \frac{1}{\mu(Q)} \right) \\ &= \sum_{q \notin Q} \left(\frac{\mu(Q \cup \{q\})}{\mu(Q)} \left(\frac{1}{\mu(Q \cup \{q\})} - \frac{1}{\mu(Q)} \right) \sum_{p \in Q} \mathcal{L}_{q,p} \right) \\ &\quad + \sum_{p \in Q} \sum_{Q' \in \mathcal{C}(Q \setminus \{p\})} \left(\frac{\mu(Q')}{\mu(Q)} \left(\frac{1}{\mu(Q')} - \frac{1}{\mu(Q)} \right) \sum_{q \notin Q} \mathcal{L}_{p,q} \right) \\ &= \sum_{q \notin Q} \frac{-\mu(q)}{\mu(Q)^2} \sum_{p \in Q} \mathcal{L}_{q,p} + \sum_{p \in Q} \sum_{Q' \in \mathcal{C}(Q \setminus \{p\})} \frac{\mu(Q) - \mu(Q')}{\mu(Q)^2} \sum_{q \notin Q} \mathcal{L}_{p,q} \\ &= \sum_{p \in Q} \frac{|\mathcal{C}(Q \setminus \{p\})| - 1}{\mu(Q)} \sum_{q \notin Q} \mathcal{L}_{p,q} \end{aligned}$$

The terms of this sum are null if $p \notin C_G$, and we deduce easily that this sum is upper bounded by a constant K which does not depend on n . By Proposition 2.8, for every $n \in \mathbb{N}$ such that $X_0^* \not\subset A_{n+1}$ almost surely, the process:

$$\begin{aligned} M_t &:= \frac{1}{\mu(X_{t \wedge \tau_n}^*)} - \frac{1}{\mu(X_0^*)} - \int_0^{t \wedge \tau_n} \mathcal{L}^* \frac{1}{\mu}(X_s^*) ds \\ &= g_n(X_{t \wedge \tau_n}^*) - g_n(X_0^*) - \int_0^{t \wedge \tau_n} \mathcal{L}^* g_n(X_s^*) ds \end{aligned}$$

is a martingale, hence:

$$\begin{aligned}\mathbb{E}\left(\frac{1}{\mu(X_{t \wedge \tau_n}^*)}\right) &= \mathbb{E}\left(\frac{1}{\mu(X_0^*)}\right) + \mathbb{E}\left(\int_0^{t \wedge \tau_n} \mathcal{L}^* \frac{1}{\mu}(X_s^*) ds\right) \\ &\leq \mathbb{E}\left(\frac{1}{\mu(X_0^*)}\right) + K\mathbb{E}(t \wedge \tau_n) \\ &\leq \mathbb{E}\left(\frac{1}{\mu(X_0^*)}\right) + Kt\end{aligned}$$

We conclude by dominated convergence that $\mathbb{E}\left(\frac{1}{\mu(X_t^*)}\right) \leq \mathbb{E}\left(\frac{1}{\mu(X_0^*)}\right) + Kt$. \square

Proof of Theorem 1.1. If condition (11) holds for every $i \in I$, then in a classical way we construct a strong stationary time from the explosion time of a strong stationary dual X^* with generator \mathcal{L}^* , and it is finite.

Conversely, let us assume that there exists $i \in I$ such that this condition does not hold. Let μ_0 be with finite support. Let X^* be a strong stationary dual of X , with generator \mathcal{L}^* . We use the separation function \mathfrak{s} , and the dominated convergence theorem together with Lemma 4.11:

$$\begin{aligned}\mathfrak{s}(t) &= \sup_{q \in G} \left(1 - \frac{\mu_t(q)}{\mu(q)}\right) \\ &= \sup_{q \in G} \mathbb{E} \left(1 - \frac{\Lambda(X_t^*, q)}{\mu(q)}\right) \\ &= 1 - \inf_{q \in G} \mathbb{E} \left(\frac{\delta_q(X_t^*)}{\mu(X_t^*)}\right) \\ &\geq 1 - \lim_{q \rightarrow \infty} \mathbb{E} \left(\frac{\delta_{\varphi_i(q)}(X_t^*)}{\mu(X_t^*)}\right) \\ &= 1 - \mathbb{E} \left(\lim_{q \rightarrow \infty} \frac{\delta_{\varphi_i(q)}(X_t^*)}{\mu(X_t^*)}\right) \\ &= 1\end{aligned}$$

Now, we know that if T is a strong stationary time of X , then for all $t > 0$, $\mathbb{P}(T > t) \geq \mathfrak{s}(t)$. So $T = \infty$ almost surely. \square

Acknowledgements: I thank my Ph.D advisor L. Miclo for introducing this problem to me and for fruitful discussions, and Pan Zhao for pointing out some imprecisions in the previous version.

References

- [1] David Aldous and Persi Diaconis. Strong uniform times and finite random walks. *Adv. in Appl. Math.*, 8(1):69–97, 1987.
- [2] William J. Anderson. *Continuous-time Markov chains*. Springer Series in Statistics: Probability and its Applications. Springer-Verlag, New York, 1991. An applications-oriented approach.
- [3] Persi Diaconis and James Allen Fill. Strong stationary times via a new form of duality. *Ann. Probab.*, 18(4):1483–1522, 1990.
- [4] Persi Diaconis and Laurent Saloff-Coste. Separation cut-offs for birth and death chains. *Ann. Appl. Probab.*, 16(4):2098–2122, 2006.
- [5] J. A. Fill and V. Lyzinski. Strong Stationary Duality for Diffusion Processes. *ArXiv e-prints*, February 2014.

- [6] James Allen Fill. Time to stationarity for a continuous-time Markov chain. *Probab. Engrg. Inform. Sci.*, 5(1):61–76, 1991.
- [7] James Allen Fill. Strong stationary duality for continuous-time Markov chains. I. Theory. *J. Theoret. Probab.*, 5(1):45–70, 1992.
- [8] James Allen Fill. An interruptible algorithm for perfect sampling via Markov chains. *Ann. Appl. Probab.*, 8(1):131–162, 1998.
- [9] James Allen Fill and Jonas Kahn. Comparison inequalities and fastest-mixing Markov chains. *Ann. Appl. Probab.*, 23(5):1778–1816, 2013.
- [10] Yu Gong, Yong-Hua Mao, and Chi Zhang. Hitting time distributions for denumerable birth and death processes. *J. Theoret. Probab.*, 25(4):950–980, 2012.
- [11] Pawel Lorek and Ryszard Szekli. Strong stationary duality for Möbius monotone Markov chains. *Queueing Syst.*, 71(1-2):79–95, 2012.
- [12] L. Miclo. Strong stationary times for one-dimensional diffusions, November 2013.
- [13] Laurent Miclo. On ergodic diffusions on continuous graphs whose centered resolvent admits a trace. *J. Math. Anal. Appl.*, 437(2):737–753, 2016.
- [14] J. R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.